(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 688 113 A2**

(12) EUROPEAN PATENT APPLICATION

(21) Application number: **95109021.6**

(51) Int. Cl.6: **H04H 5/00**

(22) Date of filing: **12.06.95**

(30) Priority: **13.06.94 JP 130653/94**

(43) Date of publication of application:
20.12.95 Bulletin 95/51

(84) Designated Contracting States:
DE FR GB

(71) Applicant: **SONY CORPORATION**
7-35, Kitashinagawa 6-chome
Shinagawa-ku
Tokyo (JP)

(72) Inventor: **Ueno, Masatoshi**
c/o Sony Corporation,
7-35, Kitashinagawa 6-chome
Shinagawa-ku,
Tokyo (JP)
Inventor: **Miyamori, Shinji**
c/o Sony Corporation,
7-35, Kitashinagawa 6-chome
Shinagawa-ku,
Tokyo (JP)

(74) Representative: **Melzer, Wolfgang, Dipl.-Ing. et al**
Patentanwälte,
Mitscherlich & Partner,
Sonnenstrasse 33
D-80331 München (DE)

(54) Method and apparatus for encoding and decoding digital audio signals and apparatus for recording digital audio

(57) An encoding method and apparatus for encoding multi-channel signals employed in , for example, a stereo system of a video disc player, a video tape recorder, a motion picture film picture system, or a so-called multi-surround acoustic system. Five channels, namely the center (C) channel, left (L) channel, right (R) channel, left surround (SL) channel and the right surround (SR) channel, for example, are handled in common depending upon frequency characteristics of digital audio signals and the targeted playback environment, and encoding (105) is done while the combinations of the channels to be handled in common are altered. High compression may be achieved with the use of pre-existing encoding and decoding units by handling the channels in common without dependency upon the degree of correlation of multi-channel digital data.

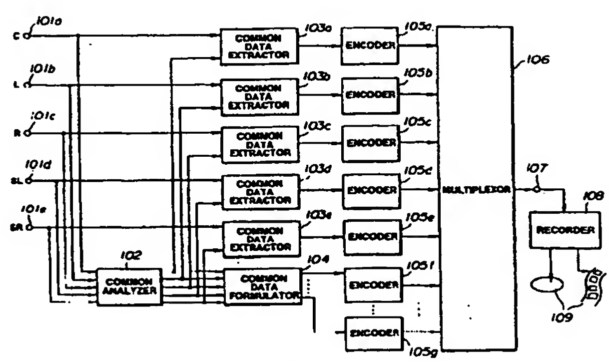

FIG.1

Rank Xerox (UK) Business Services
(3.10/3.09/3.3.4)

BACKGROUND OF THE INVENTION

This invention relates to an encoding method and apparatus for encoding multi-channel signals employed in, for example, a stereo system of a video disc player, a video tape recorder, a motion picture film picture system, or a so-called multi-surround acoustic system. The invention also relates to a corresponding decoding method and apparatus, and a recording medium.

There are a variety of techniques of high efficiency encoding of audio signals or speech signals. An example of these techniques is transform coding in which a frame of digital signals representing the audio signal on the time axis is converted by an orthogonal transform into a block of spectral coefficients representing the audio signal on the frequency axis.

There is also known a sub-band coding in which the frequency band of the audio signal is divided by a filter bank into a plurality of sub-bands without forming the signal into frames along the time axis prior to coding. In addition, there is known a combination of sub-band coding and transform coding, in which digital signals representing the audio signal are divided into a plurality of frequency ranges by sub-band coding, and transform coding is applied to each of the frequency ranges.

Among the filters for dividing a frequency spectrum into a plurality of equal-width frequency ranges include the quadrature mirror filter (QMF) as discussed in R.E. Crochiere, Digital Coding of Speech in Sub-bands, 55 Bell Syst. Tech J. No.8 (1976). With such QMF filter, the frequency spectrum of the signal is divided into two equal-width bands. With the QMF, aliasing is not produced when the frequency bands resulting from the division are subsequently combined together. In "Polyphase Quadrature Filters- A New Subband Coding Technique", Joseph H. Rothweiler ICASSP 83, Boston, there is shown a technique of dividing the frequency spectrum of the signal into equal-width frequency bands. With the present polyphase QMF, the frequency spectrum of the signals can be divided at a time into plural equal-width frequency bands.

There is also known a technique of orthogonal transform including dividing the digital input audio signal into frames of a predetermined time duration, and processing the resulting frames using a discrete Fourier transform (DFT), discrete cosine transform (DCT) and modified DCT (MDCT) for converting the signal from the time axis to the frequency axis. Discussions on MDCT may be found in J.P. Princen and A.B. Bradley, Subband Transform Coding Using Filter Bank Based on Time Domain Aliasing Cancellation", ICASSP 1987.

By quantizing the signals divided on the band basis by the filter or orthogonal transform, it becomes possible to control the band subjected to quantization noise and psychoacoustically more efficient coding may be performed by utilizing the so-called masking effects. If the signal components are normalized from band to band with the maximum value of the absolute values of the signal components, it becomes possible to effect more efficient coding.

In a technique of quantizing the spectral coefficients resulting from an orthogonal transform, it is known to use sub bands that take advantage of the psychoacoustic characteristics of the human auditory system. That is, spectral coefficients representing an audio signal on the frequency axis may be divided into a plurality of critical frequency bands. The width of the critical bands increase with increasing frequency. Normally, about 25 critical bands are used to cover the audio frequency spectrum of 0 Hz to 20 kHz. In such a quantizing system, bits are adaptively allocated among the various critical bands. For example, when applying adaptive bit allocation to the spectral coefficient data resulting from MDCT, the spectral coefficient data generated by the MDCT within each of the critical bands is quantized using an adaptively allocated number of bits.

There are presently known the following two bit allocation techniques. For example, in IEEE Transactions of Acoustics, Speech and Signal Processing, vol. ASSP-25, No.4, August 1977, bit allocation is carried out on the basis of the amplitude of the signal in each frequency band.

In the bit allocation technique described in M.A. Krassner, The Critical Band Encoder- Digital Encoding of the Perceptual Requirements of the Auditory System, ICASSP 1980, the psychoacoustic masking mechanism is used to determine a fixed bit allocation that produces the necessary signal-to-noise ratio for each frequency band.

In the high-efficiency encoding system for audio signals making use of the above-mentioned subband coding or the like, such a system in which audio data is compressed to about 1/5 by taking advantage of the characteristic of the hearing sense of the human being has been put to practice. That is, there is a system called ATRAC (Adaptive Transform Acoustic Coding, trade name by SONY Corporation) used in e.g. MiniDisc (trade name by SONY Corporation), a magneto-optical disc 64 mm in diameter, as the efficient encoding system of compressing audio data so that its data quantity becomes equal to about 1/5.

In a stereo or multi-surround audio system for a motion picture film motion picture system, high definition television, video tape recorder or a video

disc player, as well as the common audio equipments, the tendency is towards handling audio or speech signals of a plurality of, e.g., four to eight, channels. It is desired in this case to reduce the bit rate by way of high efficiency encoding.

Above all, when recording digital audio signals of eight channels, namely left channel, left center channel, center channel, right center channel, right channel, left surround channel, right surround channel and sub-woofer channel, on a motion picture film, a necessity arises for high efficiency encoding of reducing the bit rate. That is, an area sufficient to record eight channels of 16-bit linear-quantized audio data at a sampling frequency of 44.1 kHz is difficult to hold on the motion picture film, thus necessitating compression of the audio data.

The channels of the eight channel data recorded on the motion picture film are associated with a left speaker, a left center speaker, a center speaker, a right center speaker, a right speaker, a surround left speaker, a surround right speaker, and a sub-woofer speaker, which are disposed on the screen side where a picture reproduced from the picture recording area of motion picture films are projected by a projector. The center speaker is disposed at the center on the screen side, and serves to output reproduced sound by audio data of center channel. The center speaker output the most important reproduced sound, such as speech of the actor.

The sub-woofer speaker serves to output reproduced sound by audio data of sub-woofer channel. The sub-woofer speaker effectively outputs sound which feels as vibration rather than sound in low frequency range, such as sound of explosion, and is frequently used effectively in scene of explosion. The left speaker and the right speaker are disposed on left and right sides of the screen, and serve to output reproduced sound by audio data of left channel and reproduced sound by audio data of right channel, respectively. These left and right speakers exhibit stereo sound effect. The left center speaker is disposed between the left speaker and the center speaker, and the right center speaker is disposed between the center speaker and the right speaker. The left center speaker outputs reproduced sound by audio data of left channel, and the right center speaker outputs reproduced sound by audio data of right center channel. These left and right center speakers perform auxiliary roles of the left and right speakers, respectively. Above all, in movie theaters having large screen and large number of persons to be admitted, there is the drawback that localization of sound image becomes unstable in dependency upon seat positions. However, the above-mentioned left and right center speakers are added to thereby exhibit effects in creating more realistic localization of the sound image.

In addition, the surround left and right speakers are disposed so as to surround the spectator's seats. These surround left and right speakers serve to respectively output reproduced sound by audio data of surround left channel and reproduced sound by audio data of surround right channel, and have the effect to provide reverberation or impression surrounded by hand clapping or shout of joy. Thus it is possible to create sound image in more three-dimensional manner.

In addition, since defects are apt to take place on the surface of a medium of a motion picture film. If digital data is recorded as it is, data missing takes place frequently. Such a recording system cannot be employed from a practical point of view. For this reason, the ability of error correcting code is very important.

Accordingly, with respect to data compression, it is necessary to carry out compression processing to such a degree that recording can be made in the recording area on the film by taking bits for correcting code into consideration.

In this consideration, as the method of compression processing of digital audio data of eight channels as described above, there is applied the high efficiency encoding system, such as the ATRAC system, which achieves high quality comparable to CD by carrying out optimum bit allocation by taking the above-mentioned characteristics of the hearing sense into account as described above, while compressing the 16-bit digital audio data to about 1/5 with the sampling frequency of 44.1 kHz.

However, the high efficiency encoding system of compressing the digital audio data to about 1/5 is the encoding system for a single channel. If this system is employed for encoding multi-channel audio data, it is not possible to achieve effective data encoding employing data interdependency among different channels or such elements as data or format characteristics of the respective channels.

On the other hand, since the hearing sense of the human being tends to be unstable in direction feeling with respect to the sound in the high frequency range, there is known a method of encoding data in common among respective channels in the high frequency range and to record the data thus encoded in common for diminishing the recording area. However, since the level difference can be perceived, even although the direction feeling of the sound becomes indefinite, it occurs frequently that changes in the sound field be perceived by the hearer on multichannel reproduction, particularly if the correlation among different channels is low.

SUMMARY OF THE INVENTION

In view of the foregoing, it is an object of the present invention to provide a signal encoding method and apparatus, a signal decoding method and apparatus and a recording medium, in which high compression may be achieved in multi-channel signal encoding using pre-existing encoding and decoding units without dependency upon the correlation of the digital data among the respective channels.

In one aspect, the present invention provides an encoding method for encoding digital signals of plural channels and outputting the encoded digital signals and the parameter information for encoding, including the steps of handling the digital signals of at least a part of the channels in common to form a common digital signal, altering the combinations of channels handled in common depending upon frequency characteristics of the digital signals or the targeted playback environment, and encoding the common digital signal. The present invention also provides an encoding apparatus for carrying out the encoding method.

In another aspect, the present invention provides a decoding apparatus for decoding encoded digital signals using parameters for encoding, which encoded digital signals are such signals in which part or all of digital signals of plural channels are handled as one or more common signals. The combinations of channels for common handling can be altered in dependence upon frequency characteristics of the digital signals and the targeted playback environment. The decoding apparatus includes decoding means for decoding the common signals, distributing means for distributing the decoded common signals in dependence upon the combinations of common handling, and decoding means for restoring the decoded common signals of plural channels.

In still another aspect, the present invention provides a recording medium having recorded thereon such a signal in which part or all of digital signals of plural channels are handled as one or more common signals and encoded, the parameter information specifying the combinations of channels to be handled in common, an encoded signal other than the common signals and the parameter information for encoding, in addition to the parameter information concerning the encoding. The combinations of channels for common handling are altered in dependence upon frequency characteristics of the digital signals and the targeted playback environment.

With the encoding method and apparatus of the present invention, the digital signals of at least a part of plural channels are handled as common signals and encoded for raising the compression ratio. The combinations of channels to be handled in common or the processing method for handling the signals in common are altered in dependence upon the targeted or recommended playback environment for suppressing changes in the sound field otherwise caused by common handling if the digital signals are audio signals.

It is possible with the encoding method and apparatus of the present invention to evade unstable sound field due to sudden changes in the processing method of handling of common data or in the combinations of channels to be handled in common.

With the decoding apparatus of the present invention, digital signals of plural channels are decoded from at least one signal handled in common, and the processing method for handling the common signals is altered in dependence upon the recommended playback environment for the encoded signals for suppressing changes in the sound field produced by common handling if the digital signals are audio signals.

With the recording medium, such as an optical disc or a motion picture film, of the present invention, having recorded thereon the signals encoded in accordance with the encoding method and apparatus of the present invention, it becomes possible to provide a stabilized sound field.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig.1 is a schematic block circuit diagram showing a configuration of a multi-channel audio data encoding apparatus embodying the present invention.

Fig.2 is a block circuit diagram showing a configuration of a multi-channel audio data encoding apparatus not employing the present invention.

Fig.3 is a block circuit diagram showing a configuration of a common handling analyzer of the encoding apparatus embodying the present invention.

Fig.4 illustrates different sorts of selection of channels for common handling embodying the present invention.

Fig.5 illustrates changes between surround frames of the channel for common handling.

Fig.6 is a schematic block circuit diagram of a multi-channel audio signal decoding apparatus embodying the present invention.

Fig.7 is a block circuit diagram showing a configuration of a multi-channel audio decoding apparatus not employing the present invention.

Fig.8 is a block circuit diagram showing a modification of a multi-channel audio data decoding apparatus of the present invention.

Fig.9 is a block circuit diagram showing an illustrative configuration for implementing an en-

coding method of parameters for common handling embodying the present invention.

Fig.10 is a block circuit diagram showing another illustrative configuration for implementing an encoding method of parameters for common handling embodying the present invention.

Fig.11 is a block circuit diagram showing an illustrative configuration for implementing a decoding method of parameters for common handling embodying the present invention.

Fig.12 is a block circuit diagram showing another illustrative configuration for implementing a decoding method of parameters for common handling embodying the present invention.

Fig.13 illustrates different sorts of selection of channels for common handling for seven channels embodying the present invention.

Fig.14 is a block circuit diagram showing an illustrative configuration of an encoding unit of the encoding apparatus embodying the present invention.

Fig.15 is a block circuit diagram showing an illustrative configuration of a bit allocator of the encoding unit.

Fig.16 is a graph for illustrates a Bark spectrum and the masking threshold level.

Fig.17 is a graph showing a signal level, a minimum audibility curve and a masking threshold level synthesized together.

Fig.18 is a block circuit diagram showing an illustrative configuration of a decoding unit of a decoding apparatus embodying the present invention.

Fig.19 illustrates recording positions of encoded signals on a motion picture film.

Fig.20 illustrates header data of an encoded bitstream for respective channels.

Fig.21 is a diagrammatic view showing a configuration of an encoded bitstream.

Fig.22 is a block circuit diagram showing a configuration of an analyzer for common handling of another encoding apparatus of the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

Referring to the drawings, illustrative embodiments of the present invention will be explained in detail.

Fig.1 shows a configuration of an encoder (encoding apparatus) to which the encoding method of the resent invention is applied. The encoder of the present embodiment is configured to implement multi-channel encoding using a plurality of single-channel encoding units, such as encoding units of the above-mentioned encoding units of the ATRAC system.

That is, the encoder of the present embodiment is configured to encode digital audio signals of plural channels and to output the parameter information for encoding along with encoded digital audio signals. The encoder includes an analyzer for common handling 102, a common data formulator 104 and encoding units 105f to 105g as means for handling part or all of the digital audio signals of plural channels as one or plural common signals, modifying channel combinations carrying out the common handling in dependence upon frequency characteristics of the digital audio signals and the targeted playback environment and encoding the common signals.

Fig.2 shows, for comparison with the encoder of the embodiment of the present invention, a configuration of a multi-channel encoder effecting channel-based compression encoding, that is a multi-channel encoder not employing the present invention. For facilitating the understanding, similar portions of Figs.1 and 2 are represented by the same reference numerals and the corresponding description is omitted.

In the embodiment of Fig.1, explanation is made using audio data of five channels, that is a center (C) channel, a left (L) channel, a right (R) channel, a left surround (SL) channel and a right surround (SR) channel. A 5-1 channel can be constituted by adding a sub-woofer channel for ultra-low frequency. The configuration of Fig.2 is explained before explanation of Fig.1. Audio data of the center (C) channel, left (L) channel, right (R) channel, left surround (SL) channel and the right surround (SR) channel, fed via input terminals 101a to 101e, are routed to single-channel encoding units 105a to 105e, respectively. These encoding units 105a to 105e carry out encoding as later explained. The resulting encoded data are fed to a multiplexor 106 where the encoded data of the respective channels are multiplexed into a single bitstream which is outputted at an output terminal 107.

The bitstream from the output terminal 107 is recorded on a recording medium 109, such as an optical disc or a cinema film, by a recording unit 108 which effects processing such as appendage of error correction codes or modulation. Alternatively, the bitstream is transmitted over a cable or by radio transmission by a pre-set communication device.

On the other hand, with the encoder of the embodiment of the present invention, audio data of the center (C) channel, left (L) channel, right (R) channel, left surround (SL) channel and the right surround (SR) channel, supplied via the input terminals 101a to 101e, are entered to an analyzer for common handling 102. The analyzer for common handling 102 selects a technique for common han-

dling effective among the different channels, and a range to be handled in common, and selectively outputs only the portion of the audio data of the respective channels that are to be handled in common. If common handling is not performed, nothing is outputted.

Outputs of the analyzer for common handling 102 are entered to common data extractors 103a to 103e of associated channels. The common data extractors extract common portions from the original audio data from channel to channel and transmit only remaining portions to the encoders 105a to 105e. The internal configuration of the encoding units 105a to 105e is substantially equivalent to that of the encoding units of Fig.2 and hence the detailed description is omitted.

An output of the analyzer for common handling 102 is also routed to a common data formulator 104 which collects common data of the respective channels to form one or plural common data which is outputted.

On the other hand, the encoding units 105a to 105e encode outputs of the common data extractors 103a to 103e, while the encoding units 105f and 105g encode one or plural common data outputted from the common data formulator 104.

These encoding units 105a to 105 output the parameter information of common handling used for encoding, along with respective encoded data, to the multiplexor 106, which multiplexes outputs of the encoding units 105a to 105g to form a bitstream which is outputted at an output terminal 107.

The bitstream from the output terminal 107 is recorded on a recording medium 109, such as an optical disc or a motion picture film, by a recording unit 108 which effects processing such as appendage of error correction codes or modulation. Alternatively, the bitstream is transmitted over a cable or by radio transmission by a pre-set communication device.

Fig.3 shows an internal configuration of the analyzer for common handling 102.

Referring to Fig.3, the audio data of the center (C) channel, left (L) channel, right (R) channel, left surround (SL) channel and the right surround (SR) channel, supplied via the input terminals 121a to 121e, are selectively transmitted to analyzers for common handling 122a to 122i, where it is analyzed whether the use of the techniques for common handling is effective, and the results of analyses are outputted. The analyzers for common handling 122a to 122i are associated with respective selections (combinations) of the channels to be handled in common.

Outputs of the analyzers for common handling 122a to 122i are all sent to a selector for techniques of handling in common 123. The selector for techniques of handling in common 123 preferentially selects a system which allows to handle as many channels as possible, that is an output of that of the analyzers for common handling qq122a to 122i which is disposed more towards left in Fig.3. Thus the selector determines the technique for handling in common and the range of handling in common in order to output the results of selection.

The common data extractors 124 extract only the portions of data of the respective channels of the input terminals 121a to 12e, based upon data to be handled in common, obtained from the technique selector 123, and output the extracted data at associated output terminals 125a to 125e. Data on common handling, that is the parameter information for handling in common, as supplied from the technique selector 123, is outputted to data extractors for extracting commonly handled data 124a to 124e as pre-set frames on the sound frame basis.

Taking an example of five channels, namely the center (C) channel, left (L) channel, right (R) channel, left surround (SL) channel and the right surround (SR) channel, the sorts of selection of channels to be handled in common will be explained with reference to Fig.4.

Fig.4A indicates that all channels are to be handled in common. This corresponds to the analyzer 122a which analyzes effectiveness of the techniques for handling in common.

Fig.4B indicates that three channels, namely the center (C) channel, left (L) channel and the left surround (SL) channel are to be handled in common as the left-route channel, while the three channels, namely the center (C) channel, the right (R) channel and the right surround (SR) channels, are to be handled in common as right-route channels. These correspond to the analyzers for common handling 122b and the analyzers for common handling 122c, respectively.

Fig.4C indicates that two channels, namely the center (C) channel and the left surround (SL) channel, are to be handled in common as the left-route channel, while two channels, namely the right (R) channel and the right surround (SR) channels, are to be handled in common as right-route channels. These correspond to the analyzers for common handling 122d and the analyzers for common handling 122e, respectively.

Fig.4D indicates that two channels, namely the center (C) channel, left channel (L) and the right (R) channel are to be handled in common as the forward-route channel, while two channels, namely the left surround (SL) channel and the right surround (SR) channel, are to be handled in common as backward-route channels. These correspond to the analyzers for common handling 122f and the analyzers for common handling 122g, respectively.

Fig.4E indicates that two channels, namely the center (C) channel and the left channel (L), are to be handled in common as the left forward route channel, while two channels, namely the center (C) channel and the left (L) channel, are to be handled in common as right forward route channels. These correspond to the analyzers for common handling 122h and the analyzers for common handling 122i, respectively.

Thus the encoder of the embodiment illustrated allows to diminish the alien hearing feeling caused by handling in common not only by selecting the channels to be handled in common which exploits data characteristics but also by utilizing combinations of common handling which take advantage of features of targeted playback environment. Although the combinations of handling in common shown in Fig.4 are the best of those thought to be effective, any other combinations may be envisaged within the scope of the present invention.

The techniques for handling in common, shown in Fig.4, may also be changed between sound frames as pre-set frames by the relation indicated by arrows shown in Fig.5. That is, A to E in Fig.5 correspond to A to E in Fig.4, specifying that the techniques connected by arrows are changed from one sound frame to another. If, for example, the state devoid of a common channel (NONE) is selected in a sound frame, the techniques C, D or E can be selected at the next sound frame. If the technique D is selected, the technique A, E or NONE becomes selectable at the next following sound frame.

Conversely, the absence of an arrow from NONO to A or B indicates that direct selection of A or B as the next sound frame from the sound frame NONE indicating the absence of the common channel is inhibited. Sound frame localization may be improved by affording a pre-set relation in selection of the technique for common handling among sound frames.

As the technique for common handling, it is possible to cause plural techniques of common handling to exist within one sound frame.

It may be contemplated to analyze audio signals of respective channels from one specific frequency band to another and to select the technique of common handling from one specific frequency band to another. That is, such an artifice may be used in which frequency converting means for converting time-domain signals into frequency-domain signals are provided in the inside of the analyzers for common handling 122 or at an upstream side of the input terminal 121 and the resulting frequency-domain signals are analyzed from one specific frequency band to another for selecting the technique of common handling based upon the results of analyses. This enables such a processing in which common handling is realized by using the technique of common handling of all channels for the high frequency range as shown at A in Fig.4 and the technique of common handling of left and right channels for the mid frequency range as shown at C in Fig.4, thereby realizing more effective common handling.

In addition, more effective common handling may be achieved by common handling of e.g., left (L) channel and left surround (SL) channel and by common handling of e.g., right (R) channel and right surround (SR) channel within the same specified frequency range of the same sound frame.

In addition, assuming that there exist two channels each having a double power in the center channel (C), this channel may be divided and the signals of the two channels may be recorded in each of the left and left surround channels handled in common and the right and right surround channels handled in common for raising the efficiency of common handling. Furthermore, the encoding/decoding technique of separating the signal component into tonal components and noise components and encoding/decoding them as proposed by the present Assignee in our previous International Application No. PCT/JP94/00880, date of international application May 31 1994, may be used so that all channels are handled in common (A in Fig.4) for the noise components and the left and right channels are handled in common for the tonal components.

Fig.6 shows a configuration of a decoder (decoding apparatus) for carrying out the decoding method of the present invention. The decoder of the embodiment illustrated is such a decoder in which multi-channel decoding is implemented using plural single-channel decoding units, such as the decoding units corresponding to the above-mentioned ATRAC system.

The decoder of the embodiment illustrated is such a decoder in which part or all of the digital audio signals of plural channels are handled as signals of one or more channels handled in common, and in which encoded digital audio signals of plural channels, including signals for which the combinations of the channels handled in common have been changed responsive to the targeted playback environment and frequency characteristics of the digital audio signals, are decoded using the parameter information of common handling used for encoding, as shown in Fig.6. The decoder includes decoding units 133f and 133g, a distributor of data for handling in common 134 and synthesizers of data handled in common 135a to 135e. The decoding units decode the signals handled in common, distribute the signals decoded and handled in common among plural channels responsive to the parameter information for common handling

and synthesize the signals with the signals of respective channels decoded but not handled in common.

Fig.7 shows, for comparison with the decoder of the present embodiment, a configuration of a multi-channel decoder which effects channel-based decoding, that is a multi-channel decoder not employing the present invention. The portions corresponding to those of Figs.1 and 2 are denoted by the same reference numerals and the description therefor is omitted for ease of understanding.

Referring to Fig.7, before explanation of Fig.6, the encoded bitstream, entered at an input terminal 131, are demultiplexed by a demultiplexor 132 into encoded data of respective channels, which are routed to decoding units 133a to 133e. The data is decoded by decoding units 133a to 133e as later explained to decoded audio data which is outputted at output terminals 136a to 136e.

On the other hand, a bitstream entering the input terminal 131 of the decoder of the embodiment of Fig.6 is supplied to the demultiplexor 132. Since the bitstream contains data specifying the channels handled in common (parameter information of handling in common) along with channel-based encoded data and data handled in common, the demultiplexor 132 divides the encoded data and the information on the parameters of handling in common from channel to channel and transmits resulting data to the decoding units 133a to 133g.

The decoding units 133f and 133g, associated with the channels of the data handled in common, output decoded data handled in common and decoded parameter information for common handling to the distributor for data handled in common 134. The distributor for data handled in common 134 formulate data for respective channels from one or more common-handled data, with the aid of the information on the parameters for handling in common, and distributes the data to respective channels.

The synthesizers 135a to 135e of common-handled data 135a to 135e synthesize outputs of the channel-based decoding units 133a to 133e and an output of the distributor of data handled in common 134 and output the resulting data at output terminals 136a to 136e as decoded data of respective channels.

Thus the decoder of the embodiment illustrated formulates data of respective channels from one or plural data handled in common, based upon the information of the common-handling parameters, using the distributor of data handled in common 134, and synthesizes the data with data of respective channels not handled in common, using the synthesizers of common-handled data 135a to 135e for decoding digital signals of plural channels. Decoding may be achieved by the sole distributor of common-handled data 134 taking charge of distribution even if there exist plural sorts of channels made up of common-handled data or plural methods of handling un common or even if data of a specific channel is divided into plural sorts of common data which are then encoded in plural channels handled in common.

Fig.8 shows a configuration of a decoder in which common handling of multi-channel data does not depart from the single-channel encoding system employed in the embodiment illustrated and data which are not handled in common in the encoded state may be synthesized with data handled in common.

Referring to Fig.8, the demultiplexor 132 divides the bitstream entering the input terminal 131 into channel-based data not handled in common on one hand and data handled in common and the information on the parameters for common handling, on the other hand, and transmits the channel-based data not handled in common to common-encoded data synthesizing units 138a to 138e, while transmitting data handled in common and the information on the parameters for common handling to a distributor for data handled in common 137.

The distributor for data handled in common 137 formulates channel-based data from one or plural data handled in common, using the information on the parameters for common handling, and distributes the channel-based data to the respective channels.

The common handled data, outputted by the distributor for data handled in common 137 and distributed to the respective channels, and the channel-based data not handled in common, outputted by the demultiplexor 132, are routed to associated common-encoded data synthesizers 138a to 138e. The common coded data synthesizers 138a to 138e synthesize the data supplied thereto and output them as encoded data.

The decoding units 133a to 133e of the next stage decode outputs of associated common-encoded data synthesizers 138a to 138e. The outputs of associated common-encoded data synthesizers 138a to 138e are issued as channel-based data at associated output terminals 136a to 136e.

Since the distributor of data handled in common 134 and the synthesizers of data handled in common 135a to 135e are provided upstream of the decoding unit 133 in the decoder of the embodiment illustrated, the decoder may be reduced in size.

In addition, in the encoding method and apparatus of the present invention, the following technique taking advantage of the channel reproducing environment may be utilized as a technique enabling reproduction producing an alien feeling on

the hearing sense to a lesser extent.

First, the technique of altering the method of handling common data responsive to the recommended audio data reproducing environment is explained.

That is, if the five channels shown in Fig.4A are to be handled in common, the signals of the center (C) channel, left (L) channel, right (R) channel, left surround (SL) channel and the right surround (SR) channel are level-converted at a ratio of C:L:R:SL:SR = 1.0000:0.7071:0.7071:0.5000:0.5000 and subsequently synthesized. For reproduction, the data is reproduced from a sole channel or distributed at the same ratio to all channels for achieving effective common handling.

If the center (C) channel, left (L) channel and the left surround (SL) channel on one hand and the center (C) channel, right (R) channel and the right surround (SR) channel on the other hand, as shown in Fig.4B, are to be handled in common, the ratios of C:L:SL = 0.7071:1.0000:0.7071 and C:R:SR = 0.7071:1.0000:0.7071 may be employed for effective common handling.

If the left (L) channel and the left surround (SL) channel on one hand and right (R) channel and the right surround (SR) channel on the other hand as shown in Fig.4C are to be handled in common, the ratios of L:SL = 1.0000:0.7071 and R:SR = 1.0000:0.7071 may be employed for effective common handling.

If the left (L) channel, center channel (C) and the right surround (R) channel on one hand and the left surround (SL) channel and the right surround (SR) channel on the other hand, as shown in Fig.4D, are to be handled in common, the ratios of C:L:R = 1.0000:0.7071:0.7071 and SL:SR = 0.7071:0.7071 may be employed for effective common handling.

If the left (L) channel and the center (C) channel on one hand and right (R) channel and the center (C) channel on the other hand as shown in Fig.4E are to be handled in common, the ratios of C:L = 0.7071:1.0000 and C:R = 0.7071:1.0000 may be employed for effective common handling.

The above ratios are optimum values as found by experiments conducted by the present inventors and may assume different values in future experiments.

The above-described encoding method and apparatus may also be modified so that data for restoration to the data previous to common handling will be contained in the code in addition to the information of common handling in order to enable reproduction not producing alien feeling to the hearing sense.

Figs.9 and 10 show the configurations for implementing the method of extracting parameters for reproducing data used for common handling of respective channels from data handled in common. Fig.9 shows a configuration corresponding to the portions responsible for common handling processing in Fig.1, while Fig.10 shows a configuration in which common handling parameter extractor 141 is added to the configuration of Fig.9.

Figs.11 and 12 show the configuration of implementing the method of adjusting the data handled in common using the parameter contained in the code. Fig.11 shows a configuration corresponding to the portion of Fig.6 distributing the common handled data to the respective channels and Fig.12 shows a configuration in which a common handling parameter adjustment unit 142 is added to Fig.11.

In Figs.9 to 12, the constituent elements corresponding to those shown in Figs.1 and 6 are denoted by the same reference numerals and the detailed description is omitted for simplicity.

Referring first to Fig.9, data of respective channels are supplied to input terminals 101 and 101a to 101e. The data outputted by the common handling analyzer 102 and entering a common-handled data take-out unit 103 and the common handling data formulator 104 is the data used for common handling with the number of such data being equal to the number of channels of sound source data. If the data of a given channel is not used for common handling, the information specifying that such data is not employed is transmitted. On the other hand, data outputted by the common handling data formulator 104 and sent to the demultiplexor 106 is the data handled in common, with the number of the data being equal to the number of channels handled in common. The number of the channels handled in common is varied in dependence upon the method of common handling.

On the other hand, data distributed by the common handling data distributor 124 to respective channels is sent to the common handling data synthesizer 135 and synthesized with the data encoded from channel to channel so as to be outputted as channel-based decoded data.

In the configuration of Fig.9, the common handling parameter extractor 141 may be provided to each channel, as shown in Fig.10. To the common handling parameter extractor 141 are entered data and information used for channel-based common handling and all data and information handled in common. The common handling parameter extractor 141 analyzes the dependency of the channel with respect to the data of the channels handled in common and the technique of resetting the common handling in order to find the scale parameters of the channel under consideration for each frequency band or a set of frequency bands used for encoding as a unit and a scale parameter ratio between respective frequency bands in the channel under consideration. These values are sent to the

multiplexor 106 so as to be contained as the codes.

In the decoder of Fig.11, the common handling parameter adjustment unit 142 may be provided to respective channels, as shown for the decoder of Fig.12. To the common handling parameter adjustment unit 142 are entered common handling data for a channel under consideration, outputted from the common handling data distributor 134 and the common handling parameter information for a channel under consideration outputted by the multiplexor 132. By employing these data, the common handling parameter adjustment unit 142 modifies the common handling data by exploiting the technique for canceling the common handling or dependency of the data of the channel under consideration among the channels handled in common. By this, the sound field closer to the original signal data of the channel under consideration than achievable with the decoder of Fig.1 may be reproduced.

Although this system may be employed as an independent system irrelevant to the channel reproducing environment, it is possible to formulate more effective common handling parameters by analyzing the common handling parameters by taking advantage of data and the reproducing environment since data dependency degree can be predicted if the playback environment is specified.

For enabling reproduction by the encoding method and apparatus of the embodiment illustrated which does not evoke alien feeling to the hearing sense, there is an encoding method which exploits the time change information between sound frames in the selection of channels to be handled in common or the method for common handling processing.

The sound frame means an audio data processing unit for encoding and decoding and is 512 samples for the sampling frequency of 44.1 kHz for the ATRAC system employed in the present embodiment.

Meanwhile, it is possible with the encoding method and apparatus of the present embodiment to alter the selection of channels to be handled in common or the method for common handling processing from one sound frame to another. If the selection of optimum channel or processing method is done within each sound frame, it may occur that channel or processing method selection is varied from one sound frame to another to produce an alien hearing feeling due to such variation.

Thus, in the channel selection, such alien hearing feeling may be prohibited from occurring by monitoring the transition of selection from one sound frame to another to avoid continuation of common handling and non-common handling throughout the channels or to limit variation in the selection under the stationary state in which there is little change in input data.

In the selection of the method for common handling processing, frequent switching on the sound frame basis is not advisable since the difference in sound quality due to difference in the processing method is large as compared to the case of channel selection and also since the encoder and the decoder then need to be modified in dependence upon the processing method. Thus it is advisable to effect switching in terms of several sound frames at the minimum.

A modification of the present invention is hereinafter explained.

In the previous embodiment, channel selection is made from among five channels, namely the center (C), left (L), right (R), left surround (SL) and right surround (SR) channels. Fig.13 shows selection from among seven channels, namely the above five channels plus left center (CL) and right center (CR) channels.

Fig.13A shows common handling of the totality of channels.

Fig.13B shows common handling of three channels, namely center (C), left center (CL), left (L) and left surround (SL) channels, as left route channels, and common handling of four channels, namely center (C), left center (CR), right (R) and right surround (SR) channels, as right route channels.

Fig.13C shows common handling of three channels, namely left center (CL), left (L) and left surround (SL) channels, as left route channels, and common handling of three channels, namely right center (CR), right (R) and right surround (SR) channels, as right route channels.

Fig.13D shows common handling of five channels, namely center (C), left center (CL), left (L), right center (CR) and right (R) channels, as forward route channels, and common handling of two channels, namely left center (CL) and right center (SR), as backward route channels.

Fig.13E shows common handling of three channels, namely center (C), left center (CL) and left (L), as left forward route channels, and common handling of three channels, namely center (C), right center (SR) and right (R), as right forward route channels.

Fig.13F shows common handling of two channels, namely left center (CL) and left (L), as left forward route channels, and common handling of two channels, namely right center (CR) and right (R), as right forward route channels.

In the present embodiment, suitable selection of channels for common handling may be achieved by preferential processing in which the highest priority is put on the technique of Fig.13A capable of common handling of the largest number of chan-

nels and the lowest priority is put on the technique of Fig.13F.

By exploiting the combinations of common handling of channels as described above, it becomes possible to reduce he alien feeling invoked due to common handling in the case of seven channels.

In Fig.3, a sub-woofer (SW) channels may further be annexed to provide an eight-channel system. However, the sub-woofer (SW) channel is targeted for low frequency reproduction and hence is not suited for common handling. Thus the channel many bo annexed to the system without taking part in the common channel handling.

The illustrative construction and operation of the encoding unit 105 will be explained by referring to Figs.14 to 17. Fig.14 shows a construction of an encoding unit 105 for one channel.

In Fig.14, audio data corresponding to the original data from which common-handled portions have been taken out by the common-handling data take-out unit 103, that is sampled and quantized audio data, are supplied to an input terminal 24. The signals fed to the input terminal 24 are split by a spectrum splitting filter 401 into time-domain signal components in three frequency bands, namely a low frequency band of 0 to 5.5 kHz, a mid frequency band of 5.5 kHz to 11 kHz, and a high frequency band of not lower than 11 kHz, that is 11 kHz to 22 kHz.

Of the signal components of these three frequency bands from the spectrum-splitting filter 401, those of the low frequency band, mid-frequency band and the high-frequency band are sent to MDCT circuits 402L, 402M and 402H, respectively, so as to be resolved into frequency-domain signal components. The time block length for MDCT may be varied from one frequency band to another, such that, in the signal portion where signal components are changed steeply, the time block length is reduced to raise time resolution, whereas, in the stationary signal portion, the time block length is increased for effective transmission of signal components and for controlling the quantization noise.

The time block length is determined by a block size evaluator 403. That is, the signal components of the three frequency bands from the spectrum-splitting filter 401 are also sent to the block size evaluator 403 which then determines the time block length for MDCT and transmits the information specifying the thus set time block length to the MDCT circuits 402L, 404M and 402H.

Of two time block lengths for varying the time block lengths for MDCT, the longer time block length is termed a long mode and corresponds to the time duration of 11.6 msec. The short block length is termed a short mode and raises the time resolution up to 1.45 ms and to 2.9 ms for the low range of up to 5.5 kHz and for the mid range of from 5.5 to 11 kHz, respectively.

The audio signals thus resolved into signal components on two-dimensional time-frequency areas, termed block floating units, are divided by normalization circuits 404L, 404M and 404H into a sum total of 52 block floating units in the low range, mid range and in the high range, while being normalized from one block floating unit to another by way of setting scale factors.

The bit allocator 405 analyzes, by exploiting the psychoacoustic characteristics of the human auditory system, of which components the audio signals are constituted. The results of analyses are sent to a re-quantizer 406 also fed with unit-based signals from the normalization circuits 404L to 404H.

The re-quantizer 406 finds, based upon the results of analyses, the quantization steps for re-quantization of the respective units, and formulates corresponding parameters, that is decides the word lengths, while carrying out the re-quantization.

Finally, a formatter 407 assembles unit-based parameter information data and re-quantized frequency-domain signal components into a bitstream for one channel sent to the multiplexor 106 of Fig.1 in accordance with a pre-set format. An output of the formatter 407 is issued as the bitstream at an output terminal 25.

The bitstream is recorded on a recording medium, such as an optical disc or a motion picture film, by a recorder configured for effecting error correction or modulation.

The above-described encoding operation is carried out n terms of a sound frame as a unit.

The bit allocator 405 is configured as shown specifically in Fig.15.

Referring to Fig.15, the frequency-domain signal components, called hereinafter as data, are sent to an input terminal 521 from the MDCT circuits 402L, 402M and 402H.

The frequency-domain spectral data is transmitted to a band-based energy calculating circuit 522 in which the energies of the critical bands are found by calculating the sum total of the squared amplitudes of the spectral components in the respective bands. The amplitude peak values or mean values may also be employed in place of signal energy in the respective bands. Each spectral component indicating the sum value for each of the respective bands is indicated as Bark spectrum SB in Fig.16 as an output of the energy calculating circuit 522. In Fig.16, 12 bands B1 to B12 are shown as indicating the critical bands for simplifying the drawing.

It is noted that an operation of multiplying each spectral component SB by a pre-set weighting

function for taking into account the effects of masking is performed by way of convolution. To this end, an output of the band-based energy calculating circuit 522, that is each value of the Bark spectral component SB, is transmitted to a convolution filter circuit 523. The convolution filter circuit 523 is made up of a plurality of delay elements for sequentially delaying input data, a plurality of multipliers, such as 25 multipliers associated with the respective bands, for multiplying outputs of the delay elements with filter coefficients or weighting functions, and an adder for finding the sum of the outputs of the respective multipliers.

The masking means the phenomenon in which certain signals are masked by other signals and become inaudible due to psychoacoustic characteristics of the human hearing sense. The masking effect may be classified into the time-domain masking effect produced by the time-domain audio signals and concurrent masking effect produced by the frequency-domain signals. By this masking, any noise present in a masked portion becomes inaudible. In actual audio signals, the noise within the masked range is an allowable noise.

By way of a concrete example of multiplication coefficients or filter coefficients of the respective filters of the convolution filter circuit 523, if the coefficient of a multiplier M for an arbitrary band is 1, outputs of the delay elements are multiplied by coefficients 0.15, 0.0019, 0.0000086, 0.4, 0.06 and 0.007 at the multipliers M-1, M-2, M-3, M+1, M+2 and M+3, M being an arbitrary integer of from 1 to 25, by way of performing convolution of the Bark spectral components SB.

An output of the convolution filter circuit 523 is transmitted to a subtractor 524 which is employed for finding a level α corresponding to the allowable noise level in the convolved region. Meanwhile, the allowable noise level α is such a level which will give an allowable noise level for each of the critical bands by deconvolution as will be described subsequently. The subtractor 524 is supplied with an allowance function (a function representative of the masking level) for finding the level. αThe level α is controlled by increasing or decreasing the allowance function. The allowance function is supplied from a (N - ai) function generator 525 as will be explained subsequently.

That is, the level α corresponding to the allowable noise level is found from the equation:

$$\alpha = S - (n - ai)$$

where $\underline{i}$ is the number accorded sequentially to the critical bands beginning from the lower side, $\underline{n}$ and $\underline{a}$ are constants where a > 0 and S the intensity of the convolved Bark spectrum. In the equation (1), (n - ai) represents the allowance function. The values $\underline{n}$ and $\underline{a}$ may be set so that n = 38 and a = 0.5.

The level α is found in this manner and transmitted to a divider 526 for deconvolving the level α in the convolved region. By this deconvolution, the masking threshold is found from the level α. This masking threshold becomes the allowable noise level. Although the deconvolution necessitates complex arithmetic-logical steps, it is performed in the present embodiment in a simplified manner by using the divider 526.

The masking threshold signal is transmitted via a synthesizing circuit 527 to a subtractor 528 which is supplied via a delay circuit 529 with an output of the band-based energy detection circuit 22, that is the above-mentioned Bark spectral components SB. The subtractor 528 subtracts the masking threshold signal from the Bark spectral components SB so that the portions of the spectral components SB lower than the level of the masking threshold MS are masked. The delay circuit 529 is provided for delaying the signals of the Bark spectral components SB from the energy detection circuit 522 in consideration of delay produced in circuitry upstream of the synthesis circuit 527.

An output of the subtractor 528 is outputted via an allowable noise correction circuit 530 at an output terminal 531 so as to be transmitted to a ROM, not shown, in which the information concerning the number of the allocated bits is stored previously. The ROM outputs the information concerning the number of allocated bits for each band, depending on an output of the subtraction circuit 528 supplied via an allowable noise correction circuit 530.

The information concerning the number of the allocated bits thus found is transmitted to a re-quantizer 406 of Fig.14 to permit the frequency-domain data from the MDCT circuits 494L to 404H to be quantized in the re-quantizer 406 with the numbers of bits allocated to the respective bands.

In sum, the re-quantizer 406 quantizes the band-based data with the number of bits allocated in dependence upon the difference between the energy or peak values of the critical bands or sub-bands further divided from the critical bands for a higher frequency and an output of the above-mentioned level setting means.

The synthesizing circuit 527 may also be designed to synthesize the masking threshold MS and data from the minimum audibility curve RC from the minimum audibility curve generating circuit 532 representing psychoacoustic characteristics of the human hearing sense as shown in Fig.17. If the absolute noise level is lower than the minimum audibility curve RC, the noise becomes inaudible.

The minimum audibility curve differs with the difference in the playback sound level even although the encoding is made in the same manner. However, since there is no marked difference in the manner of the music entering the 16-bit dynamic range in actual digital systems, it may be presumed that, if the quantization noise of the frequency range in the vicinity of 4 kHz most perceptible to the ear is not heard, the quantization noise lower than the level of the minimum audibility curve is not heard in any other frequency range.

Thus, if the recording/ reproducing device is employed so that the noise in the vicinity of 4 kHz is not heard, and the allowable noise level is to be obtained by synthesizing the minimum audibility curve RC and the masking threshold MS, the allowable noise level may be up to the level indicated by hatched lines in Fig.17. In the present embodiment, the level of 4 kHz of the minimum audibility curve is matched to the minimum level corresponding to e.g., 20 bits. In Fig.17, the signal spectrum SS is also shown.

Besides, the allowable noise correction circuit 530 corrects the allowable noise level in the output of the subtractor 528 based on the information of the equal-loudness curve transmitted from a correction information outputting circuit 533. The equal-loudness curve is a characteristic curve concerning psychoacoustic characteristics of human hearing sense, and is obtained by finding the sound pressures of the sound at the respective frequencies heard with the same loudness as the pure tone of 1 kHz and by connecting the sound pressures by a curve. It is also known as an equal loudness sensitivity curve. The equal-loudness curve also delineates a curve which is substantially the same as the minimum audibility curve shown in Fig.17.

With the equal-loudness curve, the sound in the vicinity of 4 kHz is heard with the same loudness as the sound of 1 kHz, even although the sound pressure is decreased by 8 to 10 dB from the sound of 1 kHz. Conversely, the sound in the vicinity of 10 kHz cannot be heard with the same loudness as the sound of 1 kHz unless the sound pressure is higher by about 15 dB than that of the sound of 1 kHz. Thus it may be seen that, in the allowable noise correction circuit 530, the allowable noise level preferably has frequency characteristics represented by a curve conforming to the equal-loudness curve. Thus it may be seen that correction of the allowable noise level in consideration of the equal-loudness curve is in conformity to psychoacoustic characteristics of the human hearing sense.

Fig.18 shows an illustrative configuration of the decoding unit 133 of Fig.6 corresponding to the encoding unit 105 of Fig.1.

The decoding unit of Fig.18 decodes encoded signals for one of plural channels read by reproducing means, such as a magnetic head or an optical head, from a recording medium, such as an optical disc or a motion picture film as later explained.

In Fig.18, encoded data from the demultiplexor 132 of Fig.5 is fed to a terminal 26 and thence fed to a deformatter 411. The deformatter performs an operation which is reverse to that performed by the formatter 407, in order to produce the unit-based parameter information and the re-quantized frequency-domain signal components, that is quantized MDCT coefficients.

The unit-based quantized MDCT coefficients from the deformatter 411 are sent to a decoding circuit for the low frequency range 412L, a decoding circuit for the mid frequency range 412M and to a decoding circuit for the high frequency range 412H. These decoding circuits 412L to 412H are also fed with the parameter information from the deformatter 411. Using the parameter information, the decoding circuits 412L to 412H perform decoding and cancellation of bit allocation.

Outputs of these decoding circuits 412L to 412H are sent to associated IMDCT circuits 413L to 413H. The IMDCT circuits 413L to 413H are also fed with the parameter information and transform the frequency-domain signal components into time-domain signal components. These partial-range time-domain signal components are decoded by a band-synthesis circuit 414 to full-range signals.

An instance of recording of data encoded by the encoding method and apparatus of the present embodiment on a motion picture film as an example of the recording medium is explained by referring to Fig.19.

That is, the encoded data is recorded on a motion picture film 1 shown in Fig.19. The recording positions of the encoded data on the motion picture film 1 may be exemplified by recording regions 4 between perforations 3 of the motion picture film 1, as shown in Fig.19a, recording regions 4 between the perforations 3 on the same side of the film 1 as shown in Fig.19b, longitudinal recording regions 5 between the perforations 3 and the longitudinal edge of the film 1, as shown in Fig.19c, and by both the longitudinal recording regions 5 between the perforations 3 and the longitudinal edge of the film 1 and recording regions 4 between perforations 3, as shown in Fig.19d.

By referring to Fig.20, an instance of recording of data encoded by the encoding method and apparatus of the present embodiment on an optical film as an example of the recording medium is explained.

Fig.20 shows an example of header data of each channel, as a part of the encoded bitstream,

employed in practicing the present invention.

The header data is made up of several flags, the state of 1/0 of which specifies various conditions concerning the next following bitstream. Only part of the bitstream is disclosed herein and description on the conditions not having direct pertinence to the present invention is omitted.

The common channel handling mode is specified by a flag cplcpf. [su] and [ch] indicate the sound frame number and the channel number, respectively. The flag cplcpf is a 4-bit code and can be defined for up to a maximum of four items. If there is no "1" in any bits, that is if the flag is defined as "0000", it specifies that there is no data handled in common in the bitstream of the channel.

If the mode of handling all channels in common is selected, the flags cplcpf of all channels are set to "1000" and the all channel common handle data is entered in the first acbs (data handled in common).

If the left route channel common handle mode or the right route channel common handle mode is selected, the flag cplcpf of each channel selected for common handling of the left route channels is set to "1000", while the flag cplcpf of each channel selected for common handling of the right route channels is set to "0100". The left route channel common handling data is entered in the first acbs, while the right route channel common handling data is entered in the second acbs.

That is, which acbs is to be used for each channel may be selected by the bits of the flag cplcpf.

Thus the combinations can be varied by using the above-described encoded bitstream and the header data.

The configuration of the encoded bitstream is shown schematically in Fig.21 in which reference number 150 denotes the header of the entire bitstream, reference numbers 151 to 155 denote the data regions of each channel and reference numbers 156 to 159 denote the common-handling data regions of four channels.

The data regions 151 to 155 of each channel contains common-handling flag (cpl use flag) 160, common-handling parameters (CPL parameter) 161 and data (real data) 162. The common-handling flag (cpl use flag) 160 is made up of 4 bits (cpll-4 use bit) 170 to 173 as explained as cplchf in Fig.20.

Fig.22 shows a modification of the common handling analyzer 102.

In this figure, audio data of respective channels, that is the center (C), left (L), right (R), left surround (SL) and right surround (SR) channels, fed via an input terminal 101, are fed to orthogonal transform units 201a to 201e where they are transformed into frequency-domain signal components which are outputted.

The frequency characteristics evaluators 202a to 202e find, based upon frequency-domain signal component data for respective channels from the orthogonal transform circuits 201a to 201e, the parameters of the psychoacoustic characteristics of the human hearing sense, such as minimum audibility curve or masking threshold, and output the results along with frequency-domain signal component data.

A common-handling processing selector 203 selects, based upon the data on common handling, as obtained by evaluation by the frequency characteristics evaluators 202a to 202e and the target bit rate for encoding, such frequency range for which the absolute level of the quantization noise generated by common handling becomes lower then the minimum audibility curve. This renders the quantization noise resulting from common handling becomes inaudible. The results of selection are outputted at an output terminal 204 and thence supplied to the common handling data extractors 103a to 103e and the common handling data formulator 104. The data on common handling, outputted at the output terminal 124, is outputted in terms of a pre-set frame, such as a sound frame, as a unit.

With the above-described encoder of the present embodiment, alien hearing feeling otherwise invoked by common handling may be reduced by selecting the common handling frequency exploiting data characteristics.

The common handling technique carried out by the common handling processing selector 203 may be changed between sound frames as pre-set frames. Thus, by selecting the optimum common-handling frequency range from one sound frame to another, it becomes possible to suppress changes in the sound field otherwise produced by common handling.

It is also possible for plural selections of common handling processing in one sound frame, as shown in Fig.22. For example, if the frequency range for which the above-mentioned absolute level becomes lower than the minimum audibility curve is enhanced by independently encoding one or plural channels, without encoding in common, in a particular frequency range, common handling may be rendered more effective by effecting common handling processing in which the combinations of the channels handled in common in the pre-set frequency range are changed.

Alternatively, it is possible for plural combinations of common handling simultaneously in a particular frequency range. For example, in the case of data in which it is more effective to separately handle the forward route channels and the backward route channels (surround channels), plural

common channels having the same or different common-handled frequency ranges may be formulated for effecting common handling suited to data characteristics.

The above-described encoding and decoding methods of the present invention may be applied not only to the ATRAC system described in the embodiments but also to any other encoding system, above all, the encoding system in which time-domain signals are transformed by orthogonal transform into the frequency-domain signals.

Claims

1. An encoding method for encoding digital signals of plural channels (101a, ..., 101e) and outputting (107) the encoded digital signals and the parameter information for encoding, comprising the steps of
   handling the digital signals of at least a part of the channels in common to form a common digital signal,
   altering (103) the combinations of channels handled in common depending upon frequency characteristics of the digital signals or the targeted playback environment,
   outputting (104) the parameter information specifying the combinations of the channels handled in common,
   encoding (105) and outputting the common digital signal, and
   multiplexing (106) the parameter information and the encoded output.

2. The encoding method as claimed in claim 2, wherein the step of encoding (105) the common-handling signal includes a sub-step of adaptively altering the processing depending upon the contents of the digital signals or the advisable playback environment.

3. The encoding method as claimed in claim 2, wherein the operation of the sub-step of adaptively altering the processing is performed in terms of a pre-set time frame as a unit.

4. The encoding method as claimed in claim 1, wherein the operation of altering the combinations of channels handled in common is performed in terms of a pre-set time frame as a unit.

5. The encoding method as claimed in claim 4, wherein a plurality of the combinations of the channels handled in common are used in one frame.

6. The encoding method as claimed in any one of the preceding claims, wherein the common digital signal is a digital signal of one channel split and arrayed in at least two channels.

7. The encoding method as claimed in any one of claims 2 to 6, wherein, for at least one of the digital signals of plural channels (101a, ... 101e), the information for regenerating the pre-common-handling signal is found and the information is included in the information concerning common handling.

8. An encoding apparatus for encoding digital signals of plural channels (101a, ... 101e) and outputting the encoded digital signals and the parameter information for encoding, comprising
   means for handling the digital signals of at least a part of the channels in common to form a common digital signal,
   means (103) for altering the combinations of channels handled in common depending upon frequency characteristics of the digital signals or the targeted playback environment,
   means (104) for outputting the parameter information specifying the combinations of the channels handled in common,
   means (105) for encoding and outputting the common digital signal, and
   means for multiplexing (106) the parameter information and the encoded output.

9. The encoding apparatus as claimed in claim 8, wherein said encoding means (105) alter the processing depending upon the contents of the digital signals or the advisable playback environment.

10. The encoding apparatus as claimed in claim 8, wherein said encoding means (105) controls the alternation of the combinations of the channels handled in common from one pre-set frame to another.

11. The encoding apparatus as claimed in claim 8, wherein said encoding means (105) controls the alternation of the combinations of the processing performed on signals to be handled in common from one pre-set frame to another.

12. The encoding apparatus as claimed in claim 8, wherein said encoding means (105) split and array the digital signals of an arbitrary channel among plural channels to be handled in common.

13. The encoding apparatus as claimed in claim 8, wherein said encoding means (105) employ different sorts of combinations of channels to be handled in common in one pre-set frame.

14. The encoding apparatus as claimed in claim 8, wherein said encoding means analyzes the information for regenerating the pre-common-handling signals for part or all of the digital signals of plural channels to be handled in common and accommodates said information in the information on common handling.

15. A decoding apparatus for decoding encoded digital signals using parameters for encoding, said encoded digital signals being such signals in which part or all of digital signals of plural channels are handled as one or more common signals, with the combinations of channels for common handling being altered in dependence upon frequency characteristics of the digital signals and the targeted playback environment, comprising

decoding means (133) for decoding the common signals,

distributing means (134) for distributing the decoded common signals in dependence upon the combinations of common handling, and

decoding means for restoring the decoded common signals of plural channels based upon the signals distributed and handled in common.

16. The decoding apparatus as claimed in claim 15, wherein said decoding means (133) alter the processing depending upon the contents of the digital signals or the advisable playback environment.

17. The decoding apparatus as claimed in claim 15, wherein said decoding means (133) decode encoded signals in which plural sorts of combinations of the channels handled in common are used a plurality of number of times in the same pre-set frame.

18. The decoding apparatus as claimed in claim 15, wherein said decoding means (133) decode signals of an arbitrary channel split and arrayed in plural common signals and encoded.

19. The decoding apparatus as claimed in claim 17, wherein said decoding means (133) adjusts the signals of respective channels using the information for regenerating pre-common-handling signals contained in the information on common handling.

20. A recording medium (109) having recorded thereon such a signal in which part or all of digital signals of plural channels are handled as one or more encoded common signals, with the combinations of channels for common handling being altered in dependence upon frequency characteristics of the digital signals and the targeted playback environment, the parameter information specifying the combinations of channels to be handled in common, and an encoded signal other than the common signals and the parameter information for encoding are recorded along with the parameter information concerning the encoding.

21. The recording medium as claimed in claim 20, which is an optical disc (109) or a motion picture film (109).

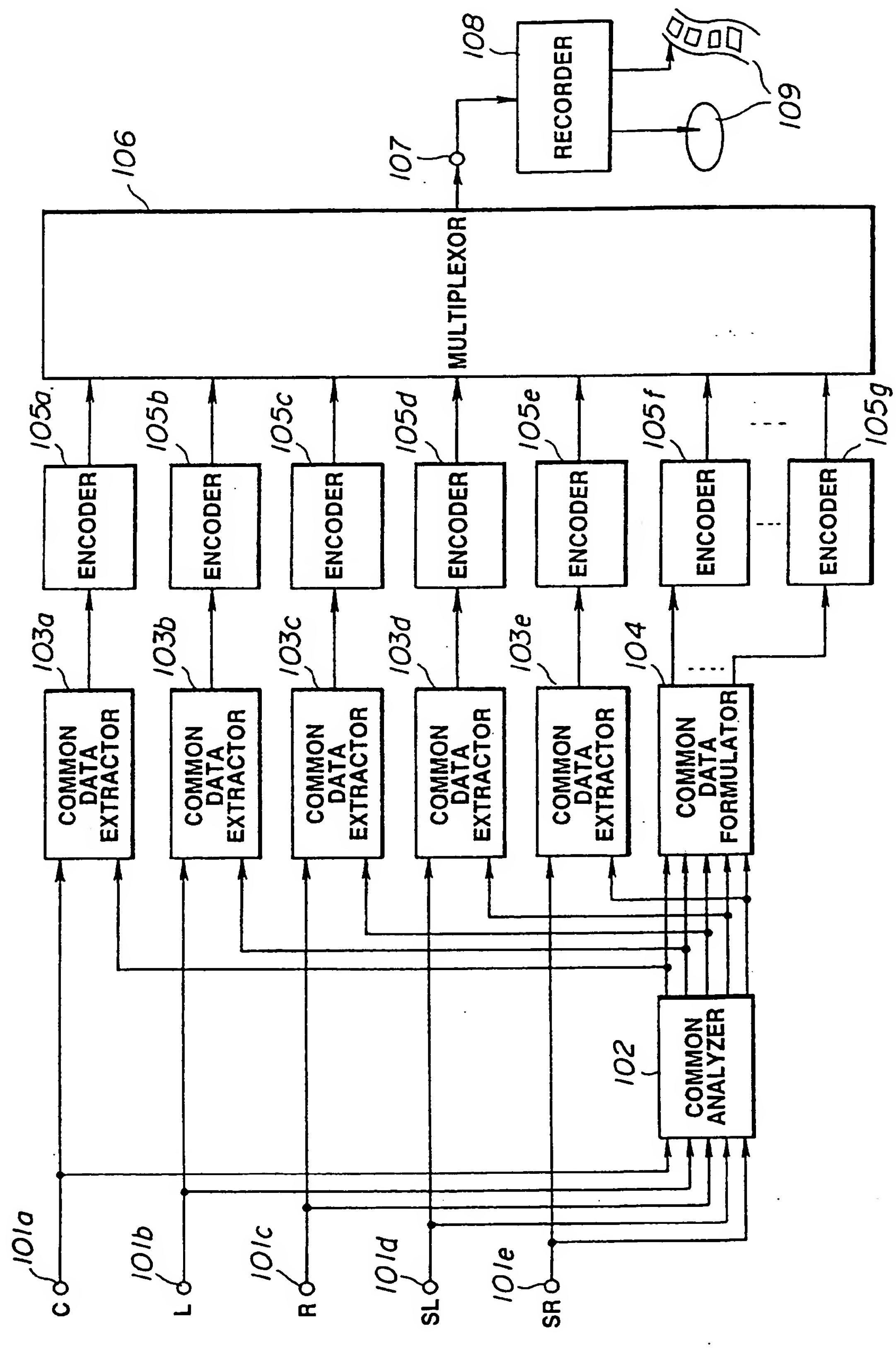
C
L
R
SL
SR
101a
101b
101c
101d
101e
COMMON ANALYZER
102
COMMON DATA EXTRACTOR
103a
103b
103c
103d
103e
COMMON DATA FORMULATOR
104
ENCODER
105a
105b
105c
105d
105e
105f
105g
MULTIPLEXOR
106
107
RECORDER
108
109

FIG.1

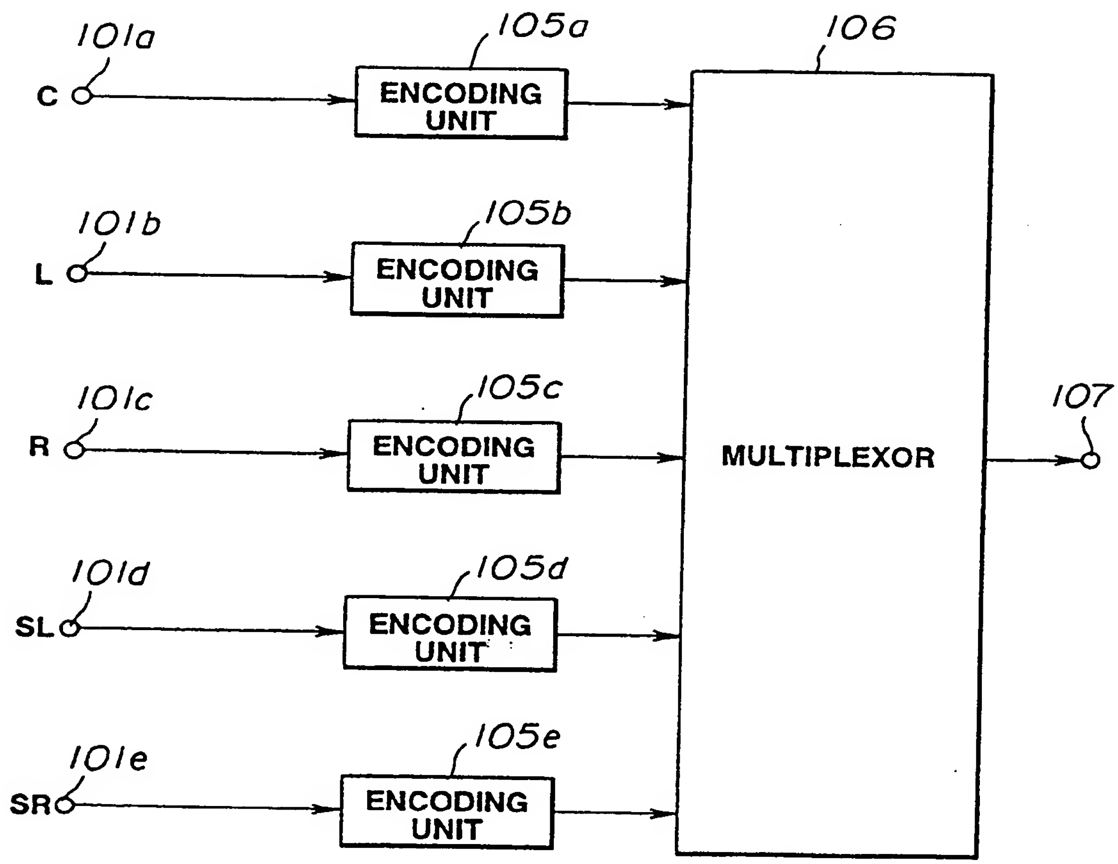

FIG.2
PRIOR ART

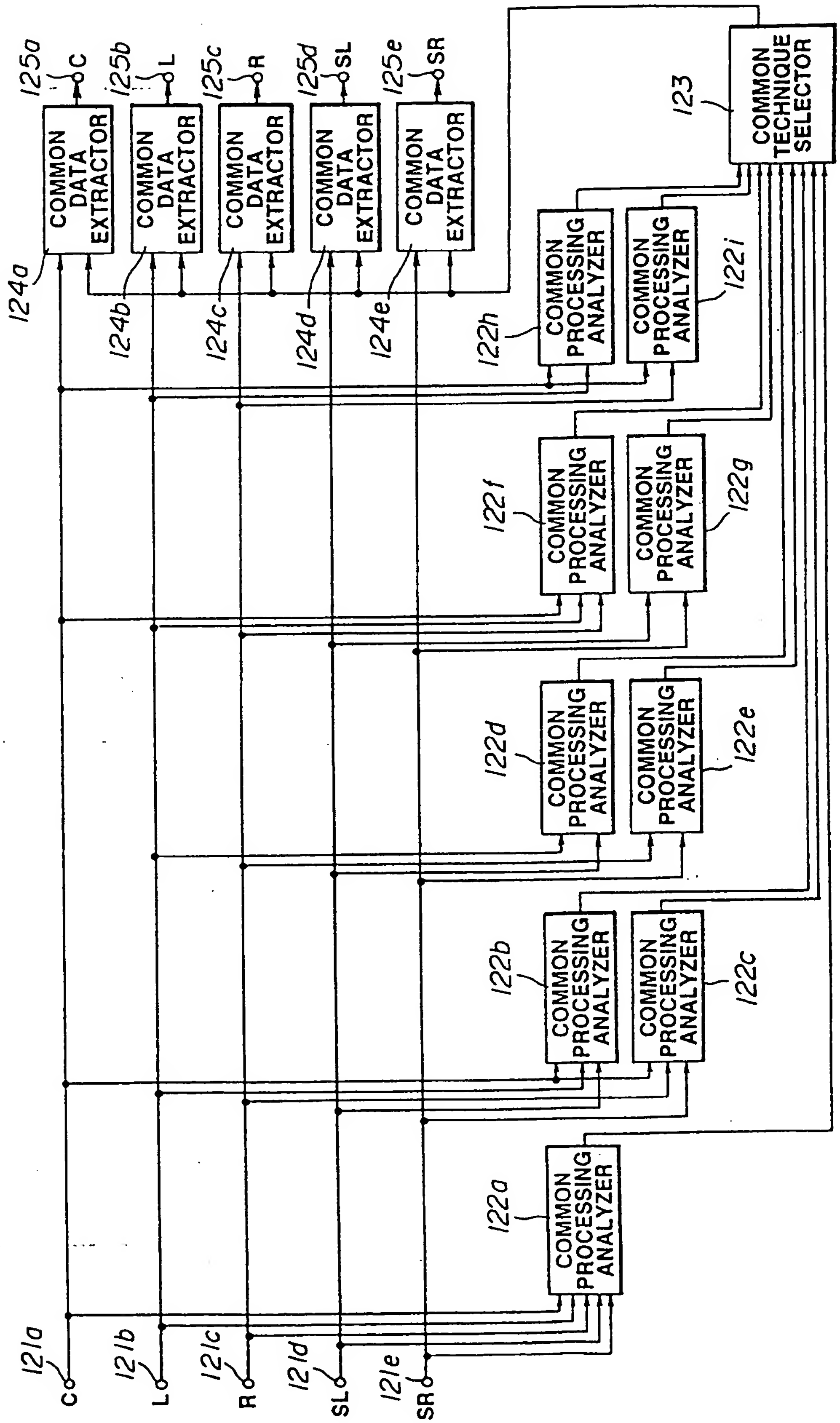
C
L
R
SL
SR
121a
121b
121c
121d
121e
122a
122b
122c
122d
122e
122f
122g
122h
122i
COMMON PROCESSING ANALYZER
123
COMMON TECHNIQUE SELECTOR
124a
124b
124c
124d
124e
COMMON DATA EXTRACTOR
125a
125b
125c
125d
125e

FIG.3

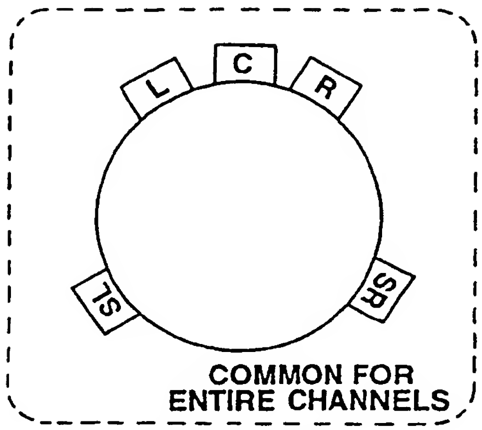

FIG.4A

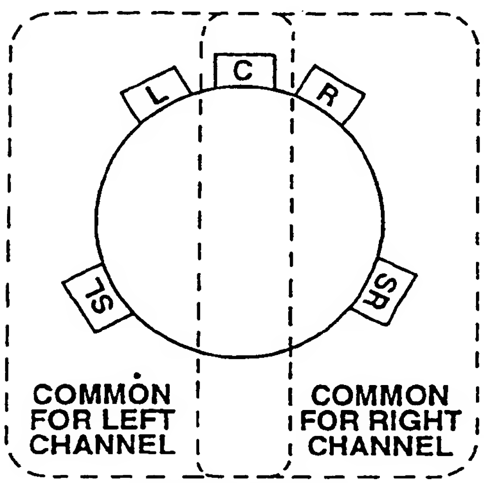

FIG.4B

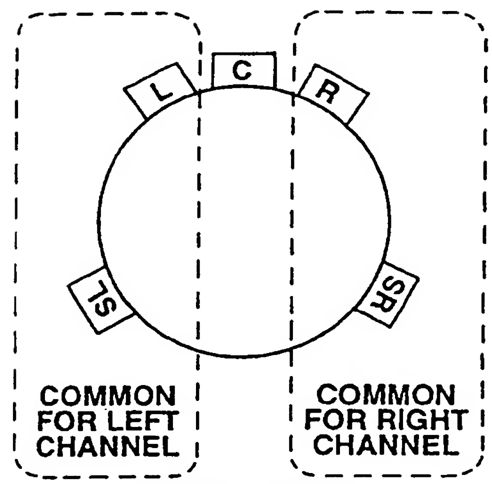

FIG.4C

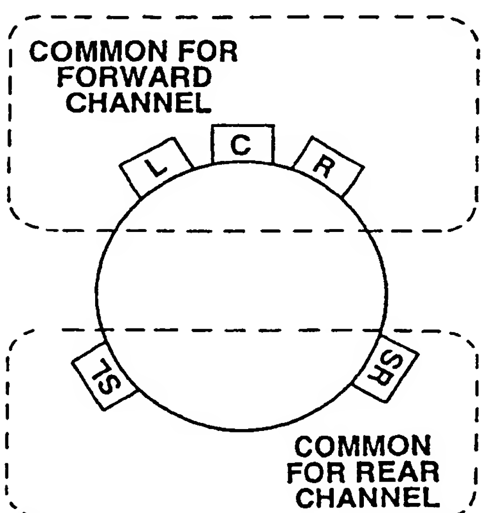

FIG.4D

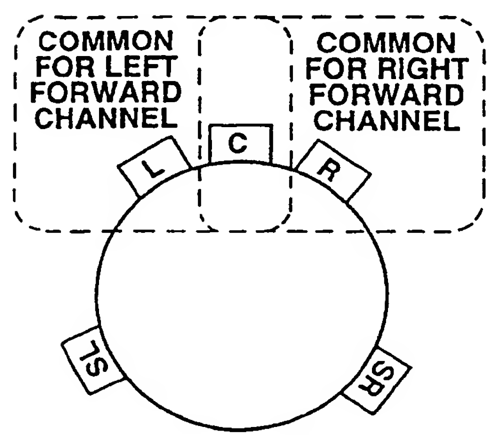

FIG.4E

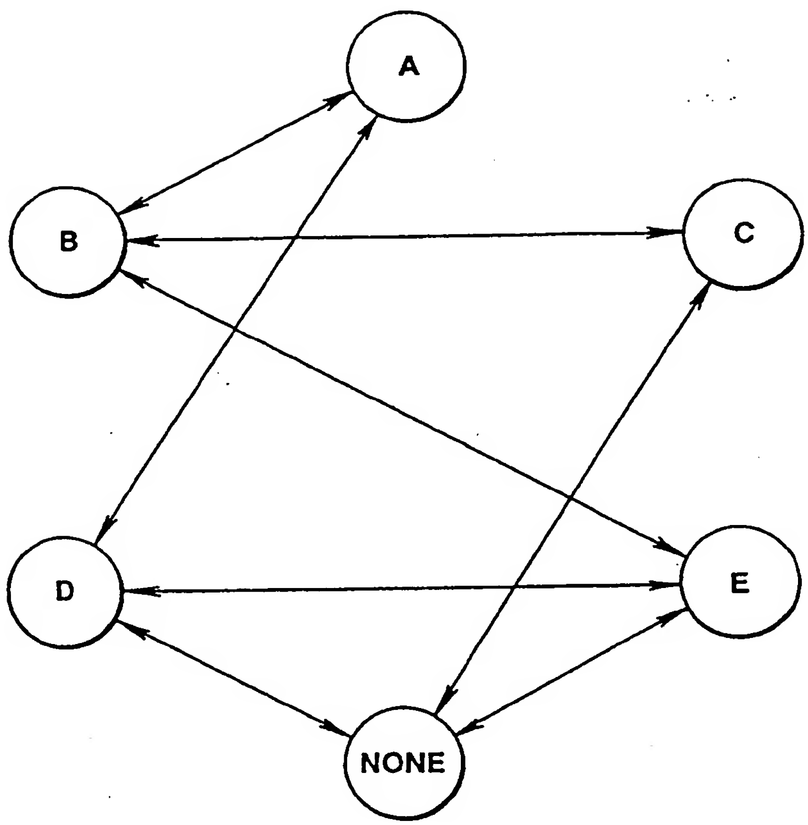

FIG.5

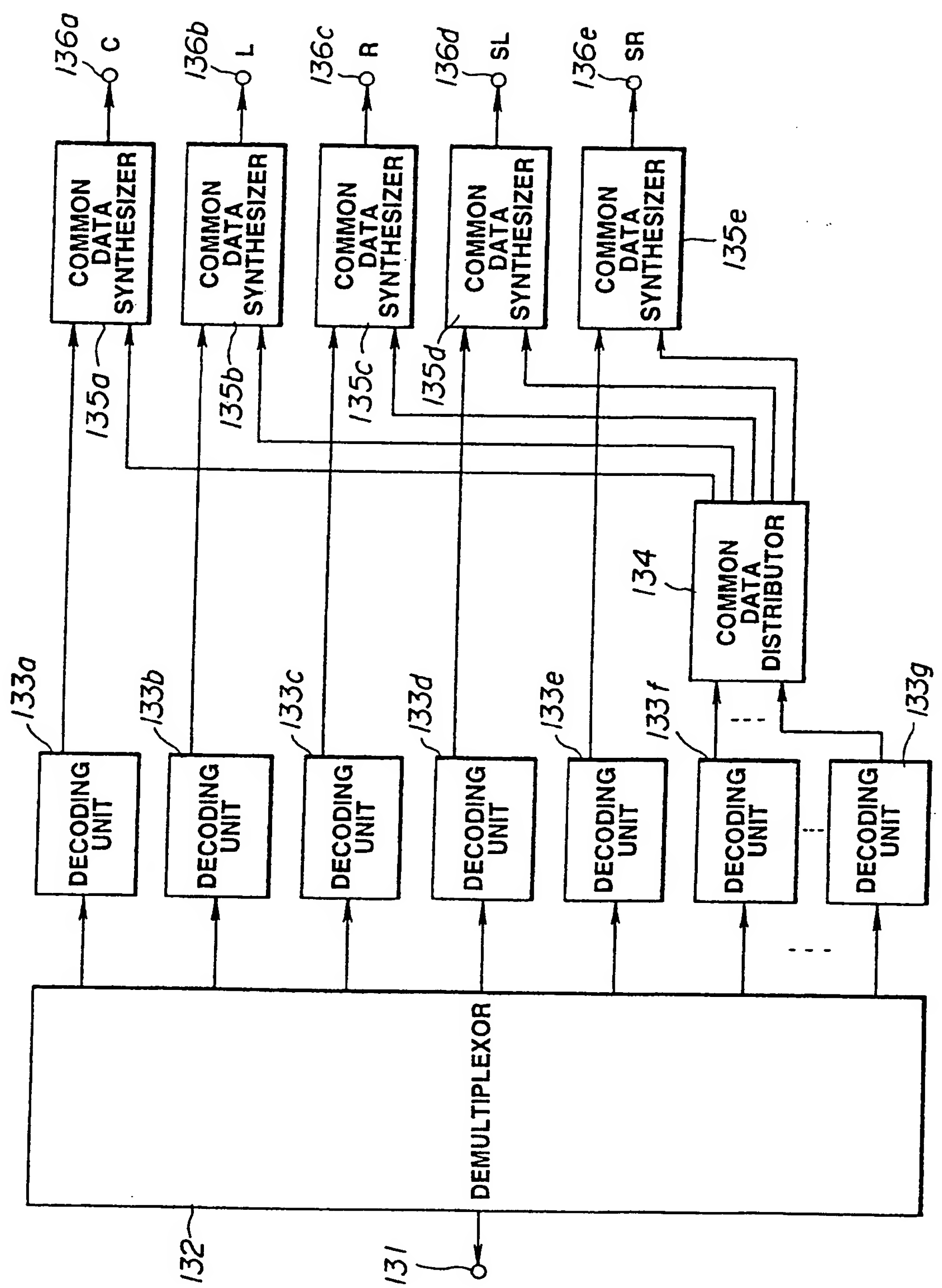
136a
C
136b
L
136c
R
136d
SL
136e
SR
COMMON DATA SYNTHESIZER
COMMON DATA SYNTHESIZER
COMMON DATA SYNTHESIZER
COMMON DATA SYNTHESIZER
COMMON DATA SYNTHESIZER
135a
135b
135c
135d
135e
134
COMMON DATA DISTRIBUTOR
133a
133b
133c
133d
133e
133f
133g
DECODING UNIT
DECODING UNIT
DECODING UNIT
DECODING UNIT
DECODING UNIT
DECODING UNIT
DECODING UNIT
DEMULTIPLEXOR
132
131

FIG.6

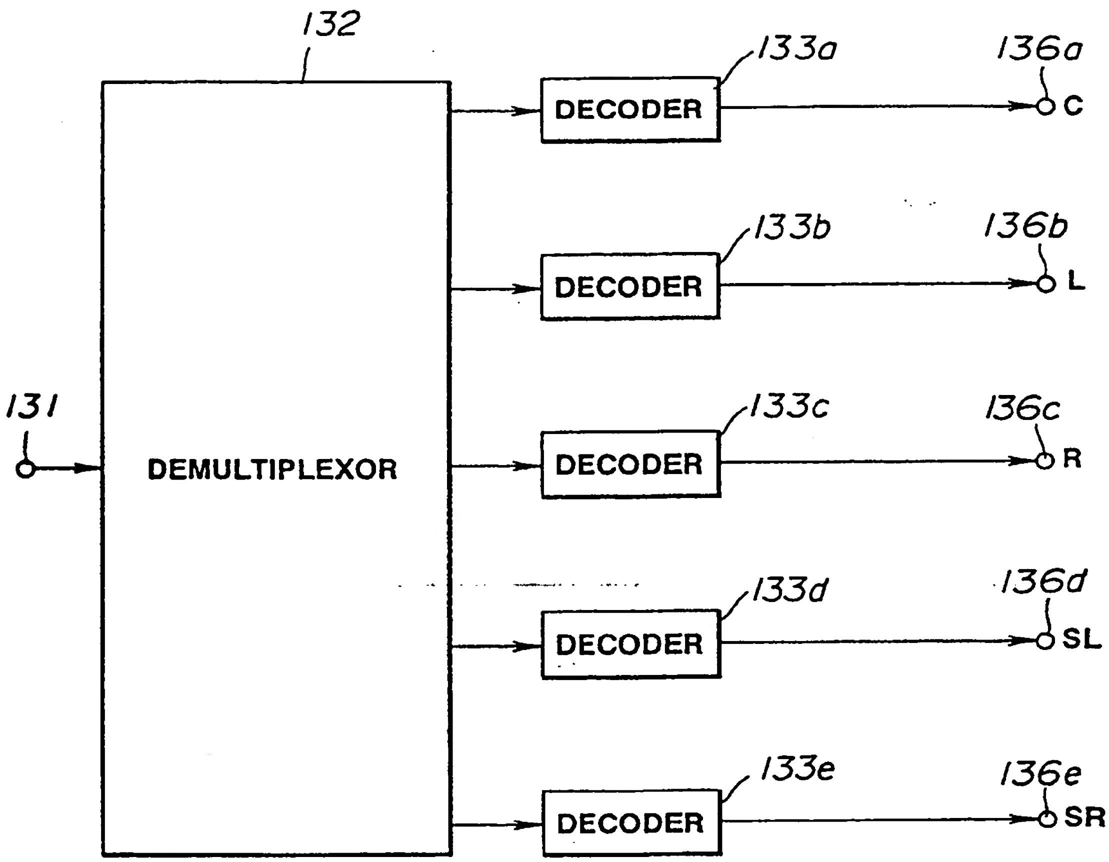
132
133a
136a
DECODER
C
133b
136b
DECODER
L
131
DEMULTIPLEXOR
133c
136c
DECODER
R
133d
136d
DECODER
SL
133e
136e
DECODER
SR

FIG.7

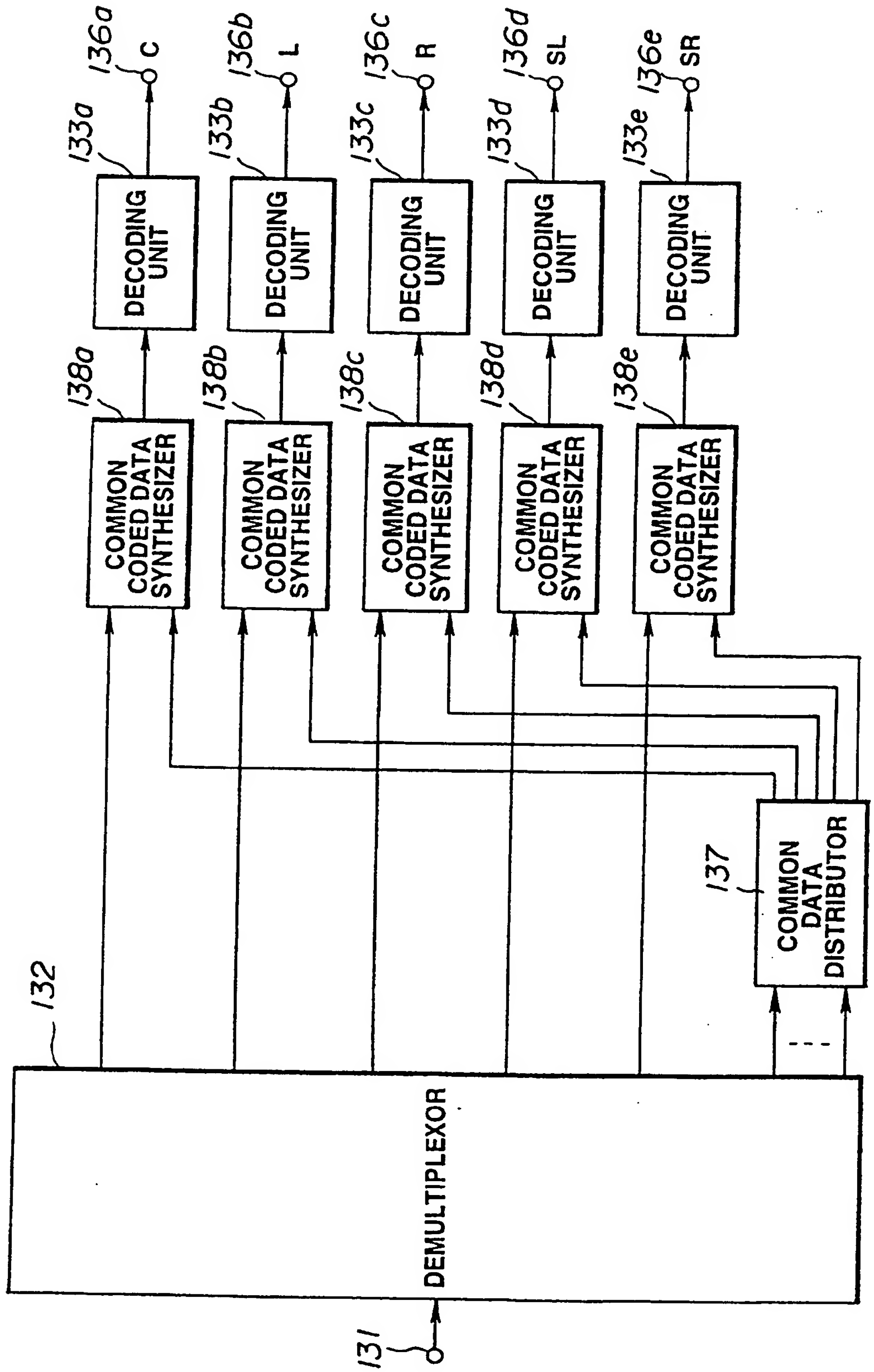
136a
C
136b
L
136c
R
136d
SL
136e
SR
133a
133b
133c
133d
133e
DECODING UNIT
DECODING UNIT
DECODING UNIT
DECODING UNIT
DECODING UNIT
138a
138b
138c
138d
138e
COMMON CODED DATA SYNTHESIZER
COMMON CODED DATA SYNTHESIZER
COMMON CODED DATA SYNTHESIZER
COMMON CODED DATA SYNTHESIZER
COMMON CODED DATA SYNTHESIZER
137
COMMON DATA DISTRIBUTOR
132
DEMULTIPLEXOR
131

FIG.8

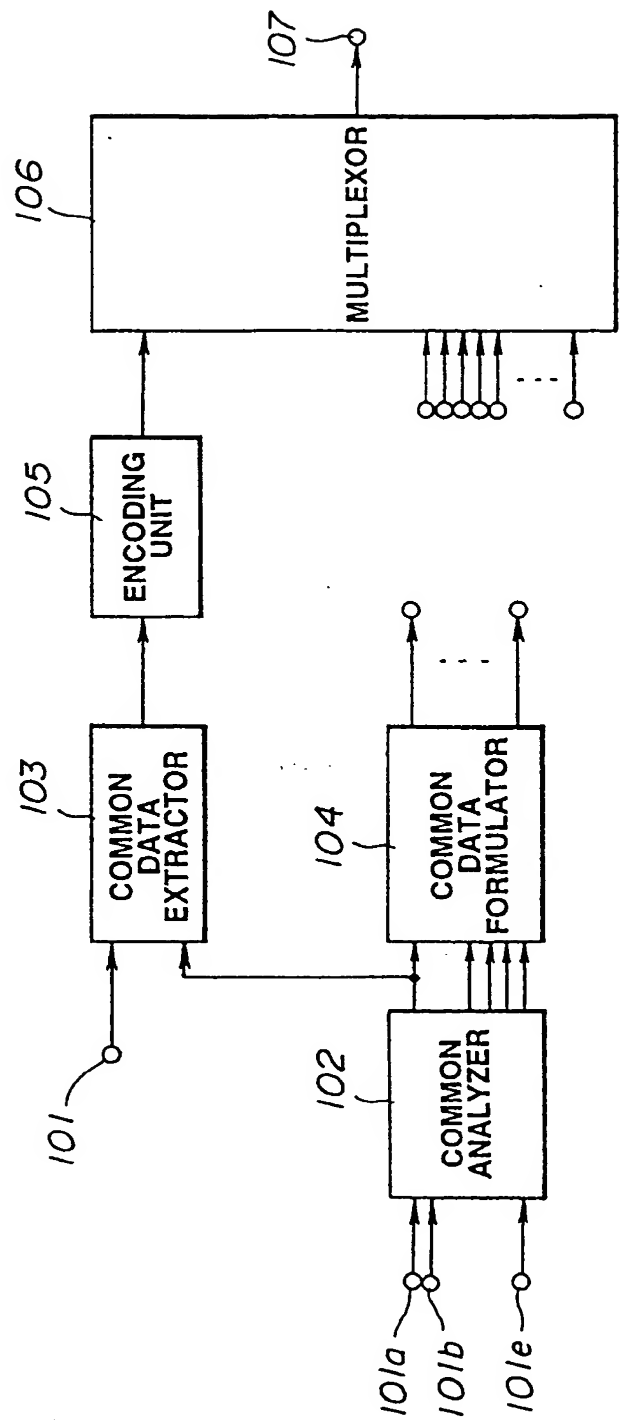
107
106
MULTIPLEXOR
105
ENCODING UNIT
103
COMMON DATA EXTRACTOR
104
COMMON DATA FORMULATOR
101
102
COMMON ANALYZER
101a
101b
101e

FIG.9

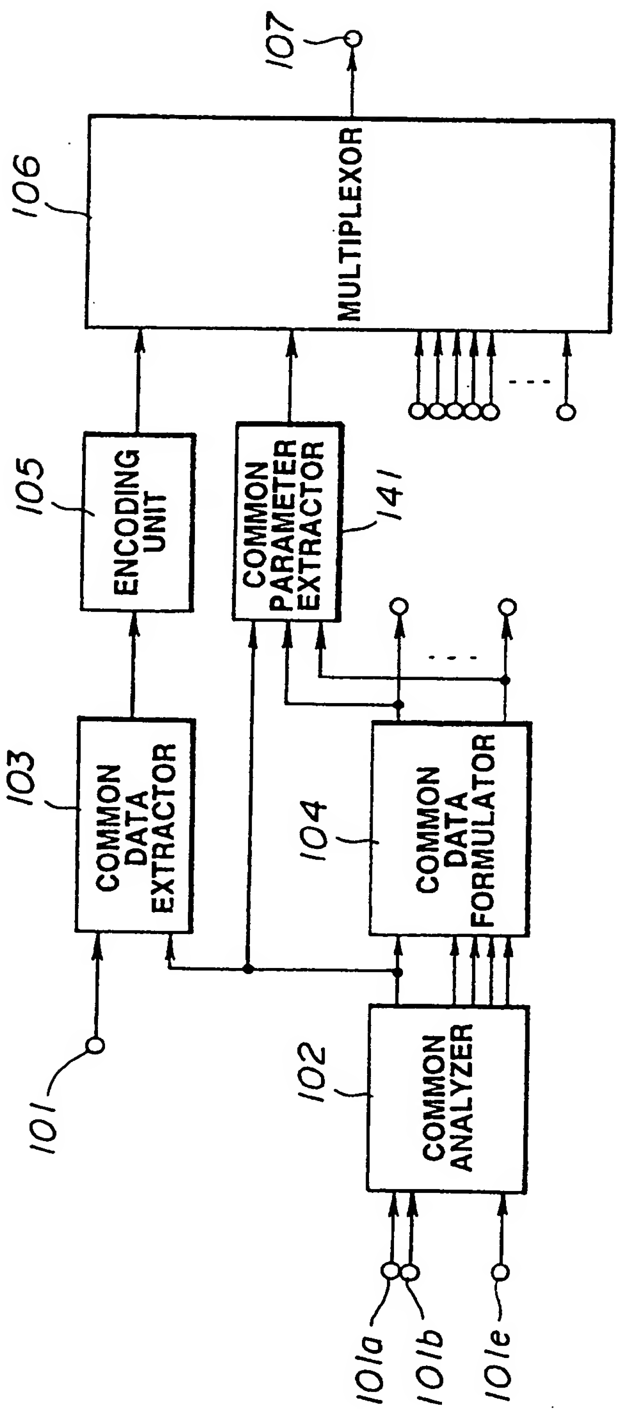
107
106
MULTIPLEXOR
105
ENCODING UNIT
COMMON PARAMETER EXTRACTOR
141
103
COMMON DATA EXTRACTOR
104
COMMON DATA FORMULATOR
101
102
COMMON ANALYZER
101a
101b
101e

FIG.10

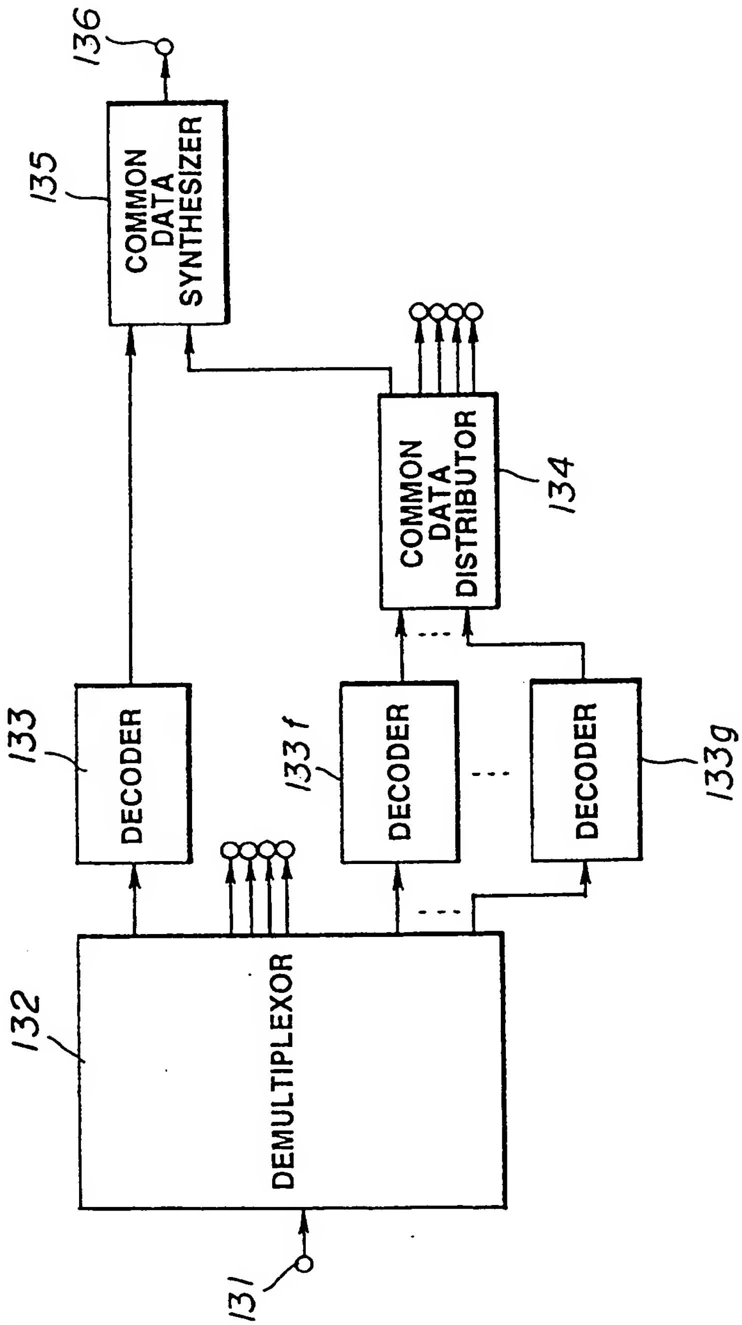
131
132
DEMULTIPLEXOR
133
DECODER
133f
DECODER
133g
DECODER
134
COMMON DATA DISTRIBUTOR
135
COMMON DATA SYNTHESIZER
136

FIG.11

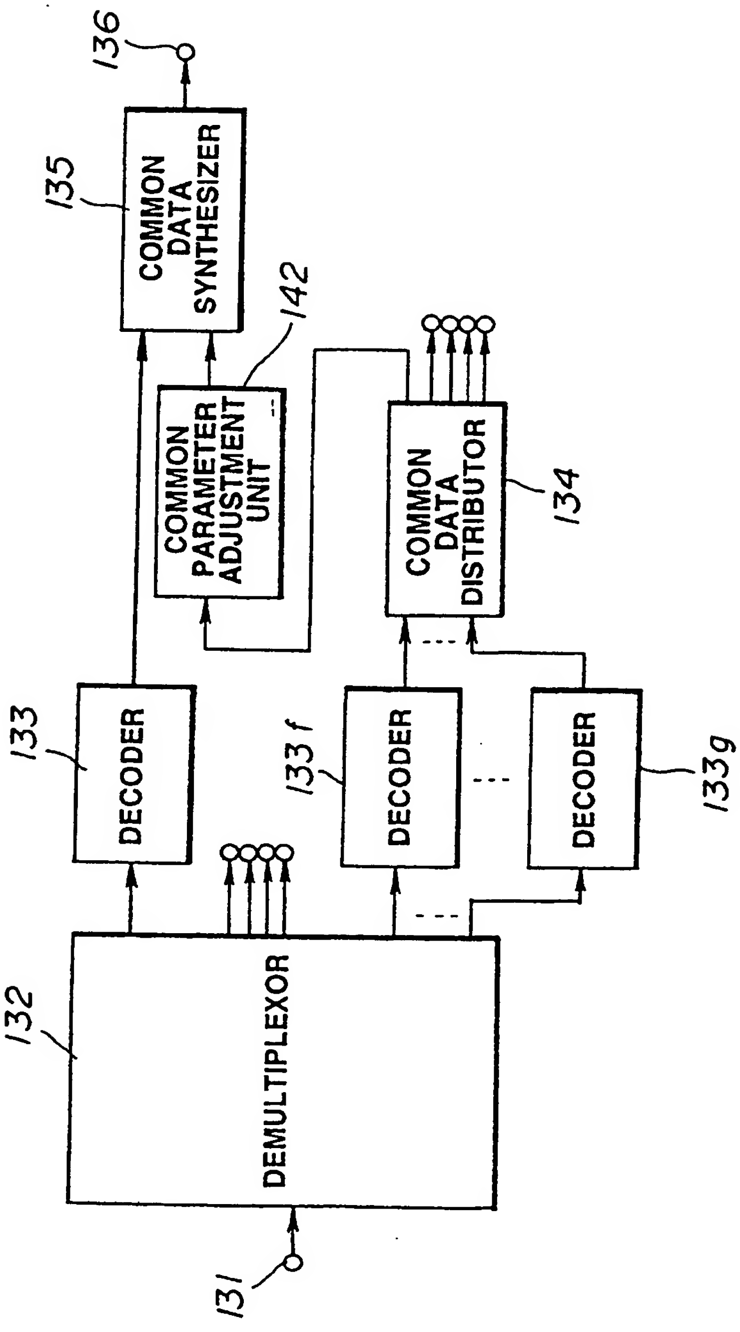
136
135
COMMON DATA SYNTHESIZER
142
COMMON PARAMETER ADJUSTMENT UNIT
COMMON DATA DISTRIBUTOR
134
133
DECODER
133f
DECODER
DECODER
133g
132
DEMULTIPLEXOR
131

FIG.12

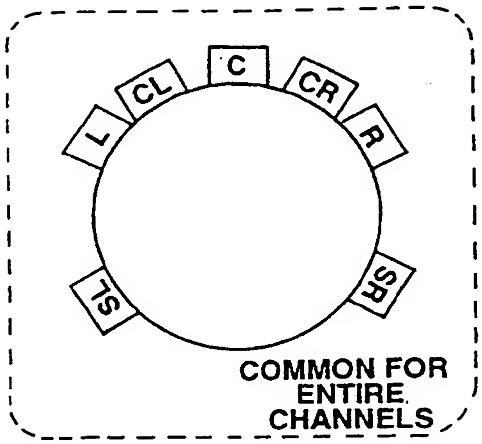

FIG.13A

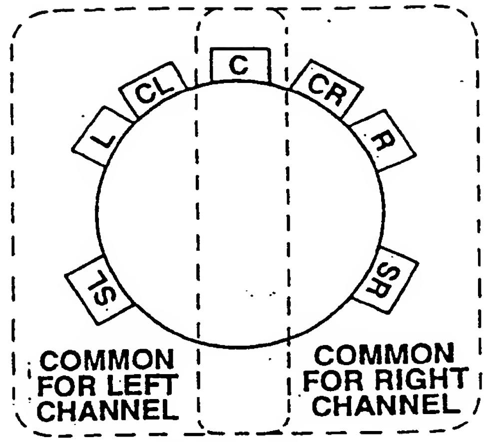

FIG.13B

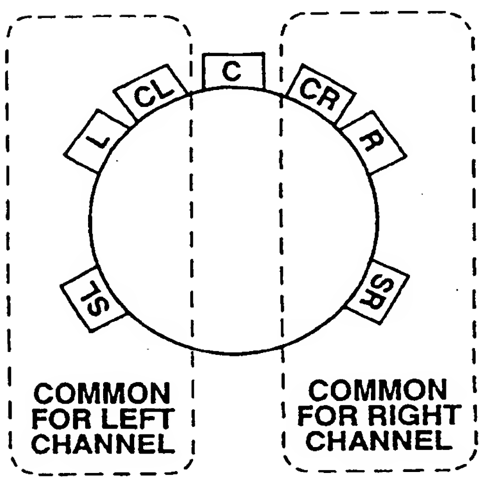

FIG.13C

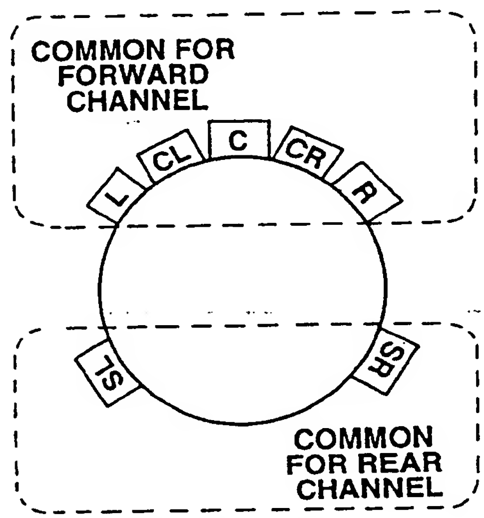

FIG.13D

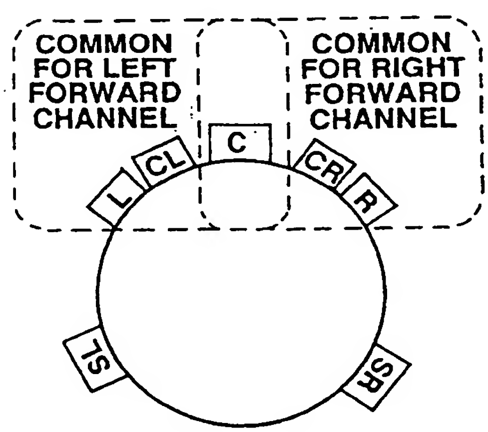

FIG.13E

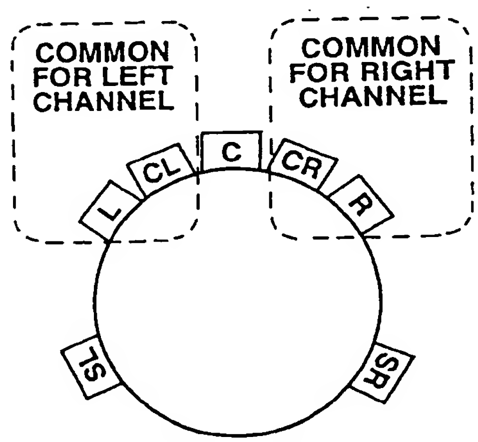

FIG.13F

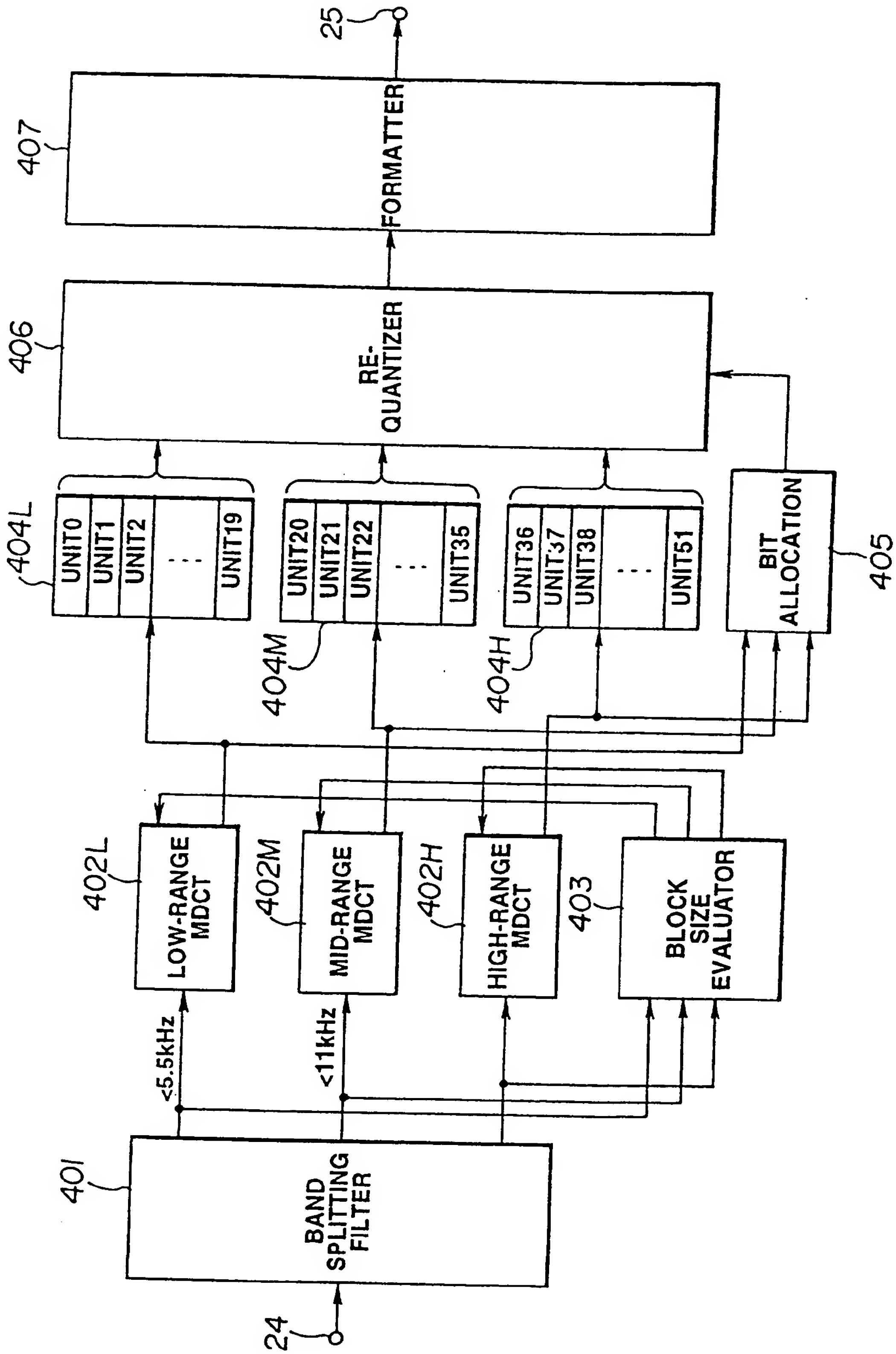
24
401
BAND SPLITTING FILTER
<5.5kHz
<11kHz
402L
LOW-RANGE MDCT
402M
MID-RANGE MDCT
402H
HIGH-RANGE MDCT
403
BLOCK SIZE EVALUATOR
404L
UNIT0
UNIT1
UNIT2
...
UNIT19
404M
UNIT20
UNIT21
UNIT22
...
UNIT35
404H
UNIT36
UNIT37
UNIT38
...
UNIT51
405
BIT ALLOCATION
406
RE-QUANTIZER
407
FORMATTER
25

FIG.14

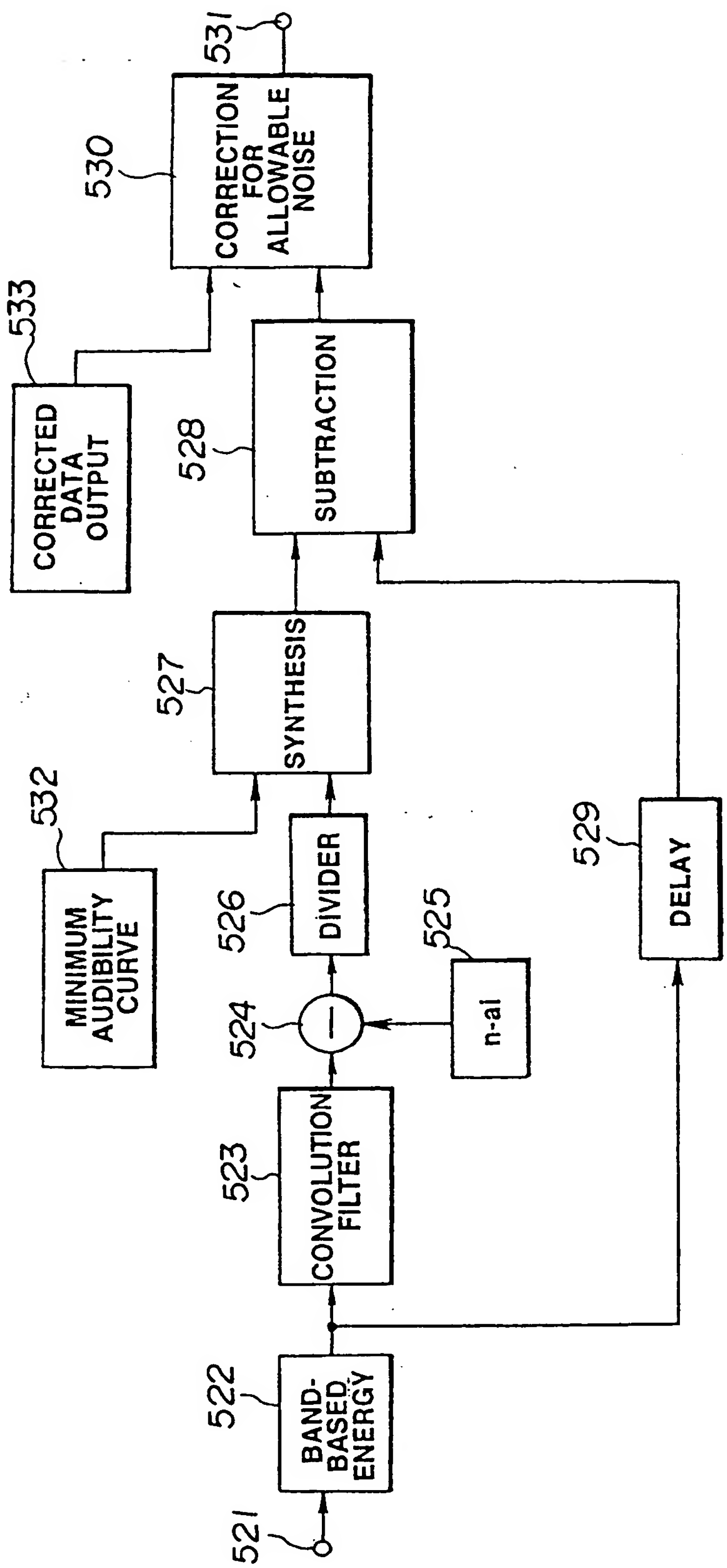
521
BAND-BASED ENERGY
522
CONVOLUTION FILTER
523
524
n-al
525
DIVIDER
526
MINIMUM AUDIBILITY CURVE
532
SYNTHESIS
527
DELAY
529
SUBTRACTION
528
CORRECTED DATA OUTPUT
533
CORRECTION FOR ALLOWABLE NOISE
530
531

FIG.15

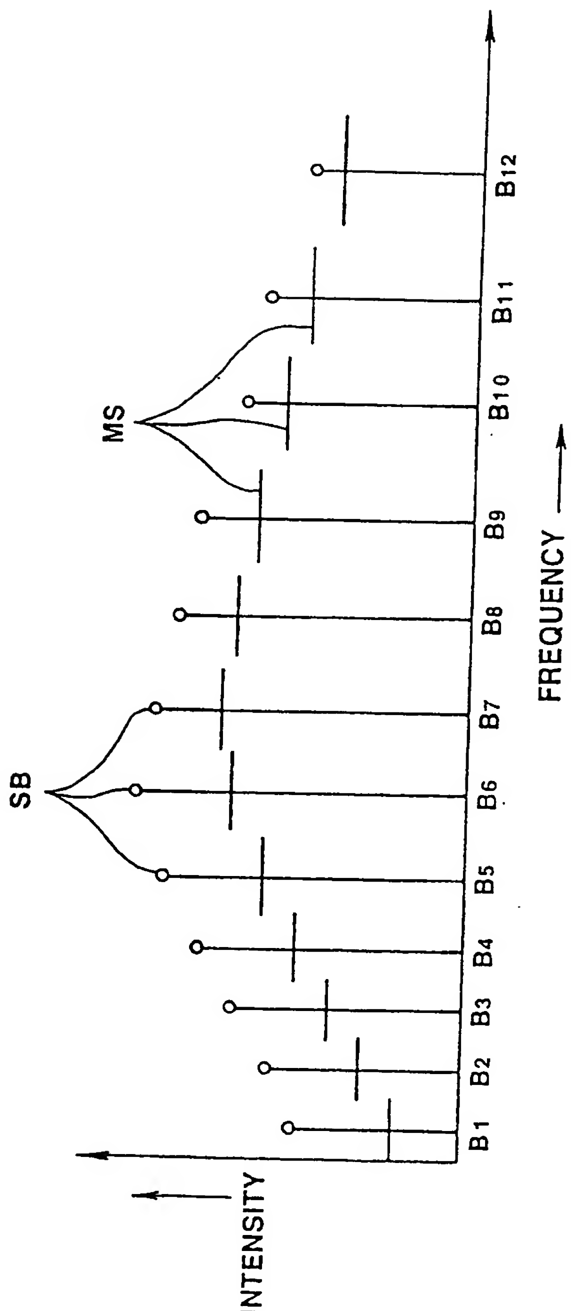

FIG.16

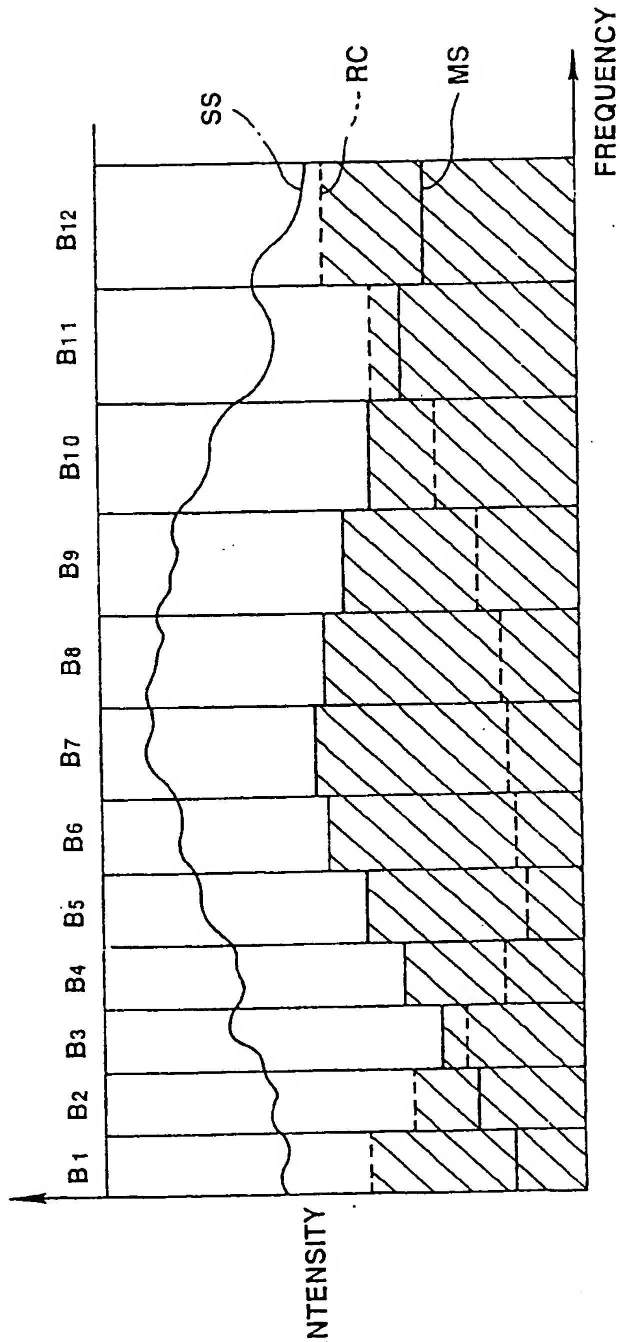
B1
B2
B3
B4
B5
B6
B7
B8
B9
B10
B11
B12
SS
RC
MS
FREQUENCY
INTENSITY

FIG.17

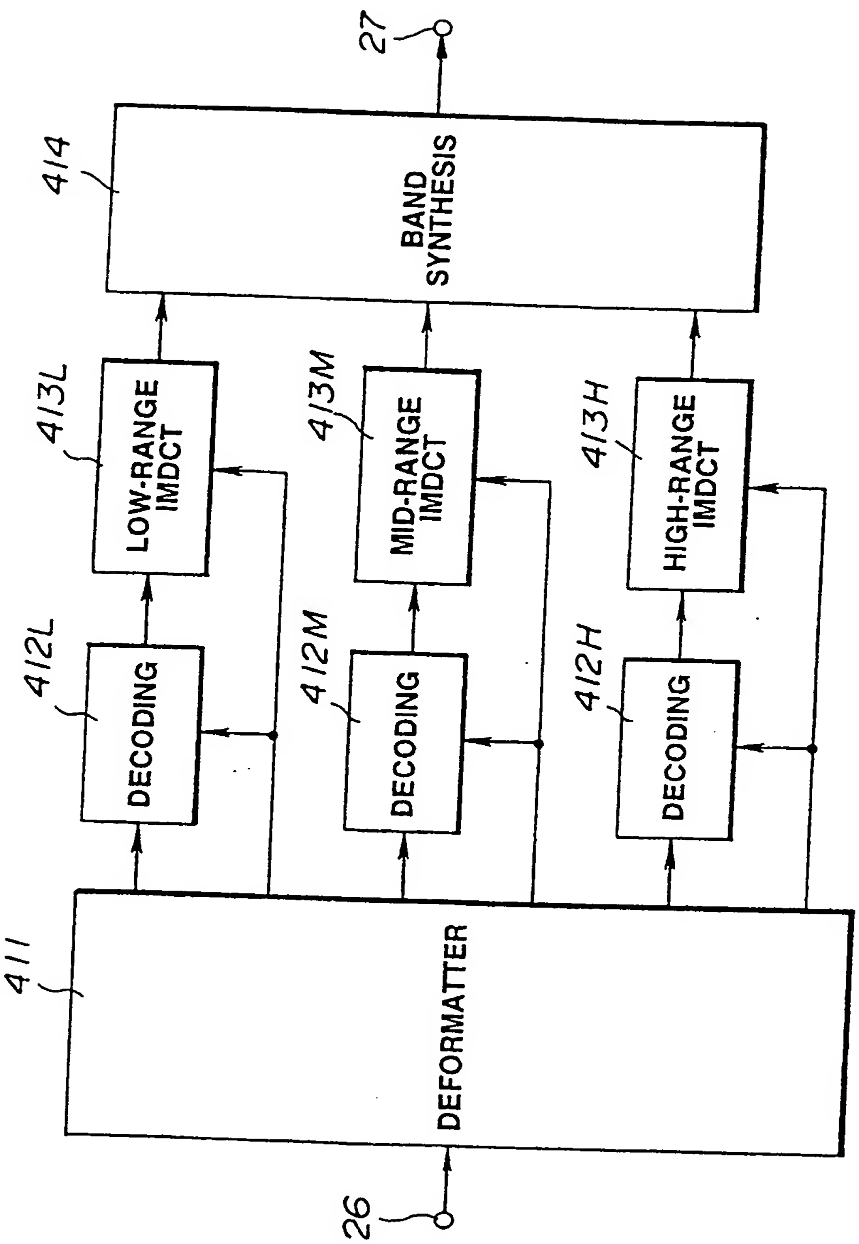
27
414
BAND
SYNTHESIS
413L
LOW-RANGE
IMDCT
413M
MID-RANGE
IMDCT
413H
HIGH-RANGE
IMDCT
412L
DECODING
412M
DECODING
412H
DECODING
411
DEFORMATTER
26

FIG.18

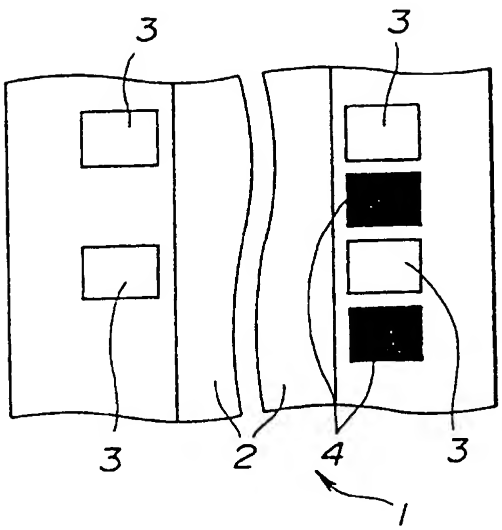

FIG.19A

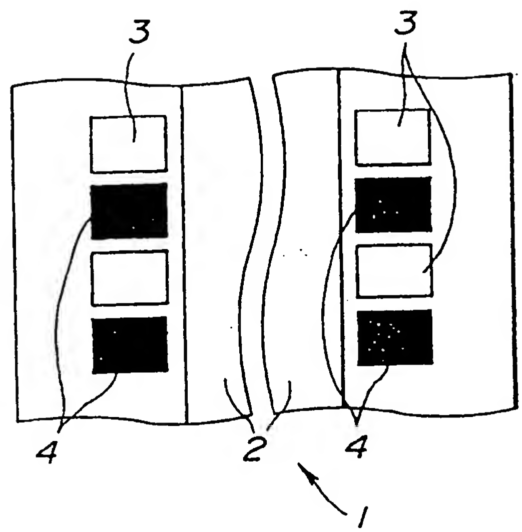

FIG.19B

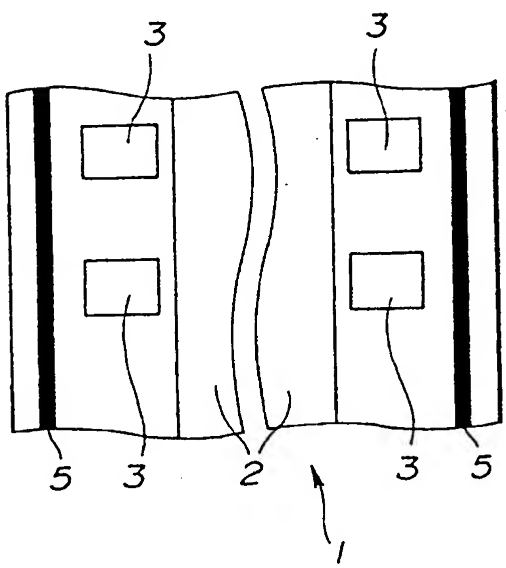

FIG.19C

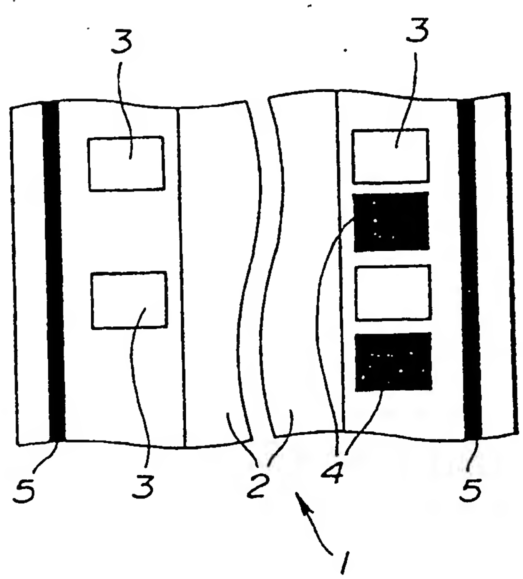

FIG.19D

```
audio_block (su)
{
    for(ch=0; ch<NChannel (acmod[suj]); ch++)
    {
        sflen [su] [ch]          LengthNum(frmsizecod[su])bits uimsf

        cplin [su] [ch]          1bit bslbf
        if (cplin [su] [ch]  ==1)
        {
            cp1_data (su, ch)
        }

        abs (su, ch)
    }

    for (cu=0;
```

$$cu< \sum_{p=0}^{3} \left(\left(\sum_{ch=0}^{NChannel(acmod[su])} cplchf[su]\,[ch]\,\delta 2^{P} \right) ?\,1:0 \right) ;$$

```
        cu++)
    {
        cpllen [su] [ch]         LengthNum(frmsizecod[su])bits uimsf
        cplst [su] [ch]          5bits uimsf
        cpled [su] [ch]          5bits uimsf
        acbs(su, ch)
        dummy = 0
    }

    if (lfeon [su]  = =  1 δ δ  0<=acmod[su] <=23)
    {
        sflenlfa [su]            8bits uimsf
        abslfe (su,ch)
    }
}


cpl-data(su,ch)
{
    cplchf [su] [ch]             4bits bslbf
    cplcost [su] [ch]            5bits uimsf
    cplcoed [su] [ch]            5bits uimsf
    if (acmod[su] = = 1 | |  28<=acmod[su] <= 31)
    {


        phsflge [su] [ch]        1bit  bslbf
        if (phsflge[su] = = 1)
        {
            for (bd=cplcost[su] [ch]; bd<=cplcoed[su] [ch] ; bd++)
            {

                phsflg[su] [ch]  [bd]   1bit bslbf
            }
        }
    }
    cplco (su, ch)
    dummy = 0
}
```

FIG.20

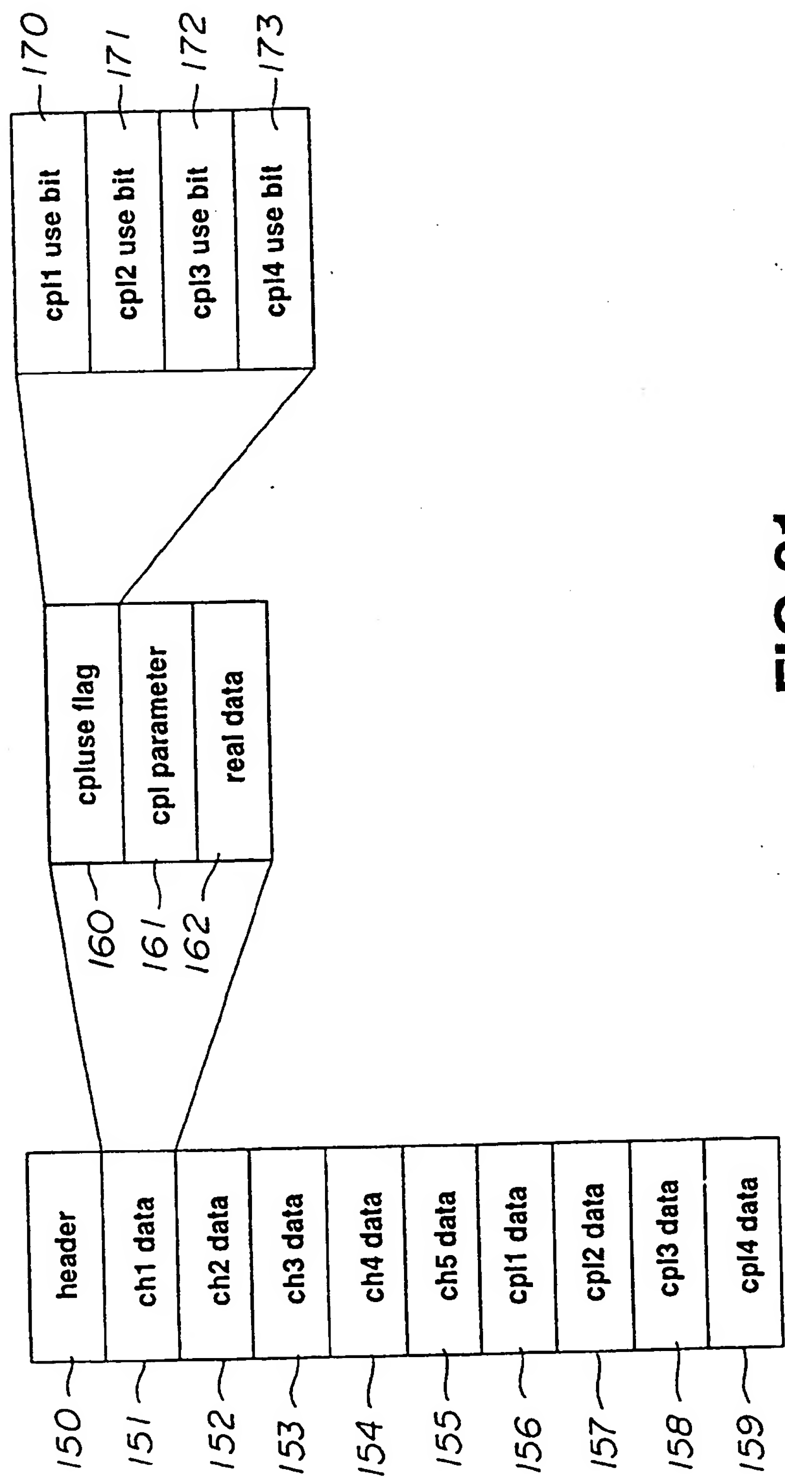
cpl1 use bit
cpl2 use bit
cpl3 use bit
cpl4 use bit
170
171
172
173
cpluse flag
cpl parameter
real data
160
161
162
header
ch1 data
ch2 data
ch3 data
ch4 data
ch5 data
cpl1 data
cpl2 data
cpl3 data
cpl4 data
150
151
152
153
154
155
156
157
158
159

FIG.21

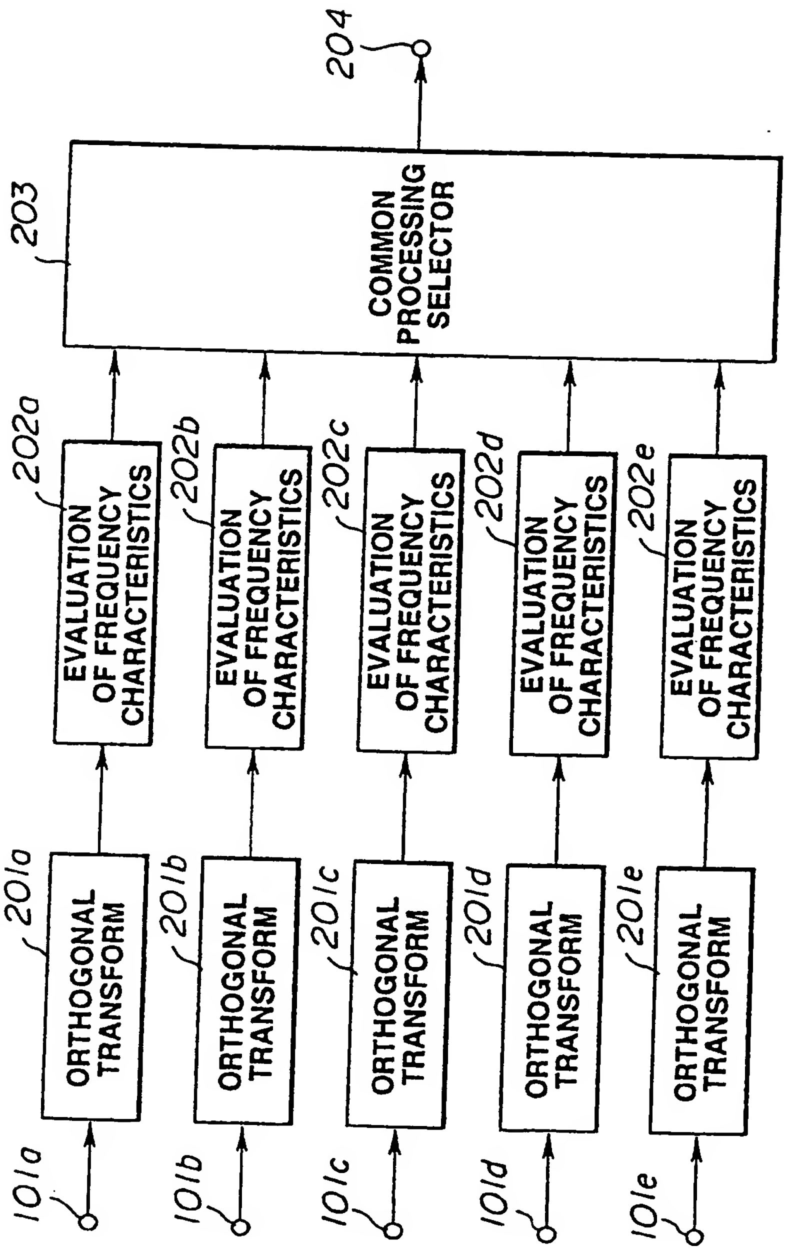

FIG.22

Docket # S&ZFH031001
Applic. # 10/679,085
Applicant: Herre et al.

Lerner and Greenberg, P.A.
Post Office Box 2480
Hollywood, FL 33022-2480
Tel: (954) 925-1100 Fax: (954) 925-1101